Observability Methods in Sensor Scheduling

by

Utku Ilkturk

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2015 by the
Graduate Supervisory Committee:

Anne Gelb, Co-Chair
Rodrigo Platte, Co-Chair
Douglas Cochran
Rosemary Renaut
Dieter Armbruster

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

Modern measurement schemes for linear dynamical systems are typically designed so that different sensors can be scheduled to be used at each time step. To determine which sensors to use, various metrics have been suggested. One possible such metric is the observability of the system. Observability is a binary condition determining whether a finite number of measurements suffice to recover the initial state. However to employ observability for sensor scheduling, the binary definition needs to be expanded so that one can measure *how observable* a system is with a particular measurement scheme, i.e. one needs a metric of observability. Most methods utilizing an observability metric are about sensor selection and not for sensor scheduling. In this dissertation we present a new approach to utilize the observability for sensor scheduling by employing the condition number of the observability matrix as the metric and using column subset selection to create an algorithm to choose which sensors to use at each time step. To this end we use a rank revealing QR factorization algorithm to select sensors. Several numerical experiments are used to demonstrate the performance of the proposed scheme.

ACKNOWLEDGEMENTS

I would like to thank Prof. Anne Gelb, Prof. Rodrigo Platte and Prof. Douglas Cochran for their unending support and guidance throughout the years. I am deeply grateful to them. I would also like to thank Dr. Brian Sadler for being instrumental in this research by introducing me to sensor scheduling and observability. I would like to express my gratitude to Prof. Rosemary Renaut, Prof. Christian Ringhofer, Prof. Dieter Armbruster and Prof. Stephen Phillips for providing me with valuable insights and suggestions on my dissertation.

I would like to thank School of Mathematical and Statistical Sciences for their support and assistance. I would like to especially thank our Graduate Program Coordinator Debbie Olson for her patience and help.

Finally, I would like to thank my family and friends for their support and encouragement.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

INTRODUCTION AND CLASSICAL OBSERVABILITY

1.1 Introduction

Most modern measurement schemes for linear dynamical systems are designed so that they have multiple sensors to select from as the system is running. Until recently, selecting sensors on the go was not feasible for the majority of systems. However, with advancements in technology, it is now often possible to schedule sensors while the linear system is propagating in time in order to use them in a more efficient way. The sensor selection/scheduling problem emerges in many fields, such as robotics [25], chemical plants [13, 52], or wireless sensor networks [44]. With the ability to schedule, the problem of trying to find the best possible sensors at each time step arises. In this regard, one needs to have a metric in order to be able to optimize the utilization of sensors.

We can formulate the sensor scheduling problem as follows. Assume that we have n potential sensors and want to select s of them at each time step with respect to some metric. Evaluating the performance of all $\binom{n}{s}$ combinations is usually not practical, and optimizing sensor selection can be shown to generally be NP-hard, [3]. In the stochastic case, where there are process and measurement noise, and the system state is estimated using Kalman filter, several different metrics and sub-optimal methods have been suggested. In this case, most common metrics used for sensor scheduling correspond to a scalar function of *estimation error covariance matrix*, [3, 44]. For example, in [27] the determinant of the estimation error covariance matrix is minimized using convex optimization. The trace of estimation error covariance matrix, which

corresponds to mean squared error, is used in [49]. This approach is well studied, [3, 9, 25, 27, 44, 49].

Another possible approach to sensor scheduling, in the noiseless case, is the *observability* of the system. If we assume there is no noise in the system and the measurements, the problem of recovering the system state becomes more similar to an inverse problem rather than an estimation problem. First defined and studied in control science by Kalman [30, 31, 32, 33], observability is a binary condition which tells us if the initial state can be determined from the measurements observed over a finite period of time. Although observability is defined for deterministic linear systems, it is still crucial in the presence of noise. If there is a component of the system state that is unobservable, then the system state cannot be deduced from the observations. Hence, observability is necessary for avoiding the ambiguity in determining the initial state as well as the state trajectory.

To be able to use observability for sensor selection, one needs to expand this binary definition of observability in order to have a *metric of observability*. Different metrics of observability for sensor selection have been proposed, see e.g. [13, 51, 52]. A typical approach, in the case of linear systems and linear measurements on the system state, is to create the observability matrices (or Gramians) corresponding to different sensor combinations and then to compare one or more criteria, e.g. singular values, trace, determinant, [51], of the resulting observability matrices to determine which sensor setup yields the optimal results for the given metric. One drawback of this approach is that it does not provide explicit guidelines for designing sensors which make the linear system observable. Rather, since all comparisons are performed a posteriori, the methods are limited to pre-determined sensor combinations. Moreover, metrics are for sensor selection rather than sensor scheduling. In other words, the observability metric seeks to find the best possible sensor placement/configuration for

designing a measurement scheme using the given sensor setups rather than the optimal sensors at each time step. To our knowledge, observability has not been utilized much for sensor scheduling problems. One of the main reasons is that in general sensor selection problems addressing observability are NP-hard, [53], and therefore it is computationally intractable to find exact solutions to large-scale problems.

This dissertation presents a new approach for using observability in sensor scheduling by considering the condition number of the observability matrix as the metric of observability, and then developing an algorithm which employs column subset selection to find best possible sensors for each time step. For this purpose, we consider two cases. First, we cover the case of time-invariant measurement schemes. We treat the observability problem with a constructive approach and present explicit methods for designing sensors to ensure an observable system. Since we cannot change sensors as time progresses in this case, we emphasize the design of sensor(s), which assure observability. Second, we study the case of time-variant measurement schemes. In this case, we incorporate a sensor schedule which employs the condition number of the observability matrix by developing a new sensor scheduling algorithm.

The document is organized as follows. This chapter gives a historical overview of observability and summarizes some classical results. Chapter 2 discusses the case of time-invariant measurement schemes and present results for designing sensors which guarantee the observability of the system. We then investigate different observability metrics and study the condition number as our metric of observability. We also present some numerical examples. Time-variant measurement schemes are discussed in Chapter 3. We first demonstrate some simple results for optimal schedules assuming the sensors can be designed without restrictions, and then consider the case where the sensors are selected from a library of possible sensors at each time step. We present a new sensor scheduling algorithm based on condition number of observ-

ability matrix, and show some more numerical examples. Our concluding remarks are presented in Chapter 4.

1.2 Classical Observability

This section provides a historical overview of observability and discusses some classical results.

Observability refers to determining the state of a linear dynamical system from the measurements over a finite time interval. In particular, in the case of discrete time linear dynamical systems, it can be reduced to determining the initial state $\mathbf{x}_0$ from a finite sequence of linear measurements. The class of discrete time linear dynamical systems of interest is described by

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \tag{1.2.1a}$$

$$\mathbf{y}_k = C\mathbf{x}_k, \tag{1.2.1b}$$

where $k = 0, 1, 2, \ldots$, the state $\mathbf{x}_k \in \mathbb{R}^n$ is the vector of system variables, the system matrix $A \in \mathbb{R}^{n\times n}$ represents the dynamics of the system, the control (input) matrix $B \in \mathbb{R}^{n\times p}$ represents possible inputs to the system, the input $\mathbf{u}_k \in \mathbb{R}^p$, the measurement matrix $C \in \mathbb{R}^{m\times n}$ represents the system sensors, and the measurement (output) $\mathbf{y}_k \in \mathbb{R}^m$.

While the first equation (1.2.1a) propagates the system in time, the second equation (1.2.1b) gives us information about the system by linearly measuring the state. We consider the case where $\mathbf{u}_k$ is identically zero. With A known, to recover the system state $\mathbf{x}_k$ from the measurements $\mathbf{y}_k$, it is sufficient to recover the initial state $\mathbf{x}_0$. In this case, determining the initial state $\mathbf{x}_0$ is equivalent to determining the

entire state trajectory $\mathbf{x}_k$, since the difference equation (1.2.1a) has a unique solution corresponding to each initial state $\mathbf{x}_0$.

To investigate observability, we first discuss classical *observability* and its dual concept *controllability*. Typically these two concepts are examined together, since one is the dual of the other in the sense that, if the system defined in (1.2.1) is controllable, then its dual system

$$\mathbf{x}_{k+1} = A^T\mathbf{x}_k + C^T\mathbf{u}_k \tag{1.2.2a}$$

$$\mathbf{y}_k = B^T\mathbf{x}_k \tag{1.2.2b}$$

is observable and vice versa. Further discussion about this duality can be found in [28].

Observability and Controllability

Introduced by Kalman in the 1960s, [30, 31, 32, 33], controllability and observability are two major concepts in control theory. These concepts roughly address the following questions, [26]:

Controllability: Does a control (or input) $\mathbf{u}_k$ always exist that can transfer the initial state $\mathbf{x}_0$ to any desired state $\mathbf{x}_k$ in finite time?

Observability: Can the initial state $\mathbf{x}_0$ of the system always be identified by observing the output $\mathbf{y}_k$ (and the input $\mathbf{u}_k$) over a finite time?

As can be seen, these concepts correspond to the relationships between the input and the state, and between the state and the output, respectively. For a linear dynamical system the answers to these questions can be characterized by the properties of the matrices A, B, C. Matrices A and B characterize controllability and hence are

called the controllability pair, and matrices A and C characterize observability, and thus called the observability pair.

Using these concepts, all linear dynamical systems can be divided into four subsystems, [22]:

1. Controllable and observable: there is a clear input-output relationship.

2. Controllable but not observable: the state can be completely controlled but some state variables (modes) cannot be determined.

3. Observable but not controllable: all the modes of the state can be determined and but there are some modes which cannot controlled.

4. Not controllable and not observable: some state modes cannot be controlled and some modes cannot be determined.

The importance of controllability and observability can be illustrated as follows. Modes that are not observable might behave in an undesired manner, however they cannot be observed. Similarly, if some uncontrollable modes act in an undesired manner, they cannot be changed by using inputs.

In this dissertation, we only consider linear dynamical systems with zero inputs ($\mathbf{u}_k = 0,\ k = 0, 1, 2, \dots$). Hence, controllability will not be discussed further in this section. In what follows, we give an algebraic condition for observability of discrete time linear dynamical systems.

1.2.1 Observability of Discrete Time Linear Dynamical Systems

Observability refers to the problem of being able to determine the initial state $\mathbf{x}_0$ of a linear dynamical system from the measurements $\mathbf{y}_k$ collected over a finite

period of time. Consider the discrete time linear dynamical system with the input $\mathbf{u}_k$ identically zero

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{1.2.3a}$$

$$\mathbf{y}_k = C\mathbf{x}_k, \tag{1.2.3b}$$

where the system state $\mathbf{x}_k \in \mathbb{R}^n$, system matrix $A \in \mathbb{R}^{n\times n}$, measurements $\mathbf{y}_k \in \mathbb{R}^m$ and measurement matrix $C \in \mathbb{R}^{m\times n}$.

Assume that we want to determine the initial state $\mathbf{x}_0$ from the measurements $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}\}$ over a finite period of time, $k = 0, 1, \ldots, t-1$ where $t \geq n$. Equations (1.2.3a) and (1.2.3b) can be rewritten as

$$\mathbf{x}_k = A^k\mathbf{x}_0,$$

$$\mathbf{y}_k = CA^k\mathbf{x}_0.$$

Hence we can write the following linear system of equations:

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{t-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t-1} \end{bmatrix} \mathbf{x}_0 = \Phi_t\mathbf{x}_0. \tag{1.2.4}$$

The initial state $\mathbf{x}_0$ can be uniquely determined if and only if Φ_t is non-singular, i.e. $\mathrm{Ker}\,(\Phi_t) = \{0\}$. If Φ_t would have a non-zero nullspace then any non-zero initial state in the nullspace cannot be distinguished from the zero initial state $\mathbf{x}_0 = 0$, since the outputs $\mathbf{y}_k$ would be all zero. Thus before giving a definition of observability, it might be useful to define what an unobservable state is.

Definition 1.2.1. [15] An initial state $\mathbf{x}_0$ is called ***unobservable*** if for any $T > 0$, $\mathbf{x}_0$ produces the output $\mathbf{y}_k = 0$, $k = 0, 1, \ldots, T$.

Definition 1.2.2. [15] The system (1.2.3) is called ***(completely) observable***, if no initial state is unobservable (except zero). If any non-zero unobservable states exist, then the system is called ***unobservable***.

Note that by Cayley-Hamilton theorem, A^n can be expressed as a linear combination of $\{A^0, A^1, \ldots, A^{n-1}\}$, [29]. Thus $\mathrm{Ker}\,(\Phi_t) = \mathrm{Ker}\,(\Phi_n)$ for $t \geq n$, which means it is enough to consider Φ_t up to time $(n-1)$. Now we can define the observability matrix:

Definition 1.2.3. The ***observability matrix*** $\Phi = \Phi(A, C)$ of the linear dynamical system (1.2.3) is defined as

$$\Phi = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}. \tag{1.2.5}$$

Theorem 1.2.1. *The system (1.2.3) is* ***observable*** *if and only if the observability matrix* Φ *in (1.2.5) is non-singular.*

Proof. Let the measurements $\mathbf{y} = [\mathbf{y}_0, \ldots, \mathbf{y}_{n-1}]^T$, then (1.2.4) becomes

$$\mathbf{y} = \Phi \mathbf{x}_0. \tag{1.2.6}$$

Thus, we can determine $\mathbf{x}_0$ uniquely if and only if Φ is non-singular. The initial state $\mathbf{x}_0$ can be recovered by

$$\mathbf{x}_0 = \left(\Phi^T \Phi\right)^{-1} \Phi^T \mathbf{y}. \tag{1.2.7}$$

□

We now provide two simple examples of an observable and an unobservable system.

Example 1.2.1. Observable System, [38]

Consider the discrete time linear system where $\mathbf{x}_k \in \mathbb{R}^3$ and

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}.$$

From (1.2.5), we have

$$\Phi = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 3 & 1 \\ 7 & 12 & 1 \end{bmatrix},$$

which is non-singular. Hence the system is completely observable.

Example 1.2.2. Unobservable System, [38]

Now consider a system with

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

In this case we have

$$\Phi = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 2 & -3 \end{bmatrix},$$

which is singular with rank two. Hence, the system is not completely observable, i.e. we cannot uniquely recover the initial condition $\mathbf{x}_0$ from the observations $\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \mathbf{y}_2 \end{bmatrix}^T$.

Now that we defined observability of a discrete time linear dynamical system, we continue with some important classical results. In particular, we present results about how an observable subsystem can always be extracted out of a system even when the system itself is not observable, and moreover how the state space $\mathbb{R}^n$ can be separated into observable and unobservable subspaces. These results in turn help us to state methods for designing measurement schemes assuring the observability of the system.

1.2.2 Separation of Observable Part

As can be seen in Example 1.2.2, not all systems are completely observable. However even if a discrete time linear dynamical system is not completely observable, we can still find a subsystem which is completely observable. In other words, even though we cannot determine all the components of $\mathbf{x}_0$ we can still recover at least its projection onto a subspace of $\mathbb{R}^n$.

Theorem 1.2.2. [26] *Consider the linear dynamical system defined in (1.2.3). If* $\text{rank}(\Phi) = q < n$*, i.e. the observability matrix defined in (1.2.5) is not full rank, then a non-singular matrix* T *exists which transforms the linear dynamical system into the following equivalent form:*

$$\begin{bmatrix} \hat{\mathbf{x}}_{k+1}^{(1)} \\ \hat{\mathbf{x}}_{k+1}^{(2)} \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & 0 \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k}^{(1)} \\ \hat{\mathbf{x}}_{k}^{(2)} \end{bmatrix}$$
$$\mathbf{y}_k = \begin{bmatrix} \hat{C}_1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k}^{(1)} \\ \hat{\mathbf{x}}_{k}^{(2)} \end{bmatrix},$$

where $\hat{\mathbf{x}}_k^{(1)} \in \mathbb{R}^q$, $\hat{\mathbf{x}}_k^{(2)} \in \mathbb{R}^{n-q}$, *with the submatrices* $\hat{A}_{11}$, $\hat{A}_{21}$, $\hat{A}_{22}$, $\hat{C}_1$ *of compatible size. Moreover, the subsystem*

$$\hat{\mathbf{x}}_{k+1}^{(1)} = \hat{A}_{11}\hat{\mathbf{x}}_k^{(1)}$$
$$\mathbf{y}_k = \hat{C}_1\hat{\mathbf{x}}_k^{(1)},$$

is completely observable. (Note that $\hat{\mathbf{x}}_0^{(1)} \neq \mathbf{x}_0$*)*

Proof. A suitable transformation matrix T can be constructed in the following way. Let

$$T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix},$$

where T_1 consists of linearly independent rows of the observability matrix Φ and T_2 contains $(n-q)$ arbitrary rows such that T becomes non-singular. Then by defining $\mathbf{x}_k = T\hat{\mathbf{x}}_k$, we have

$$\mathbf{x}_{k+1} = A\mathbf{x}_k$$
$$T^{-1}\mathbf{x}_{k+1} = T^{-1}A\mathbf{x}_k$$
$$\hat{\mathbf{x}}_{k+1} = T^{-1}AT\hat{\mathbf{x}}_k$$
$$\hat{\mathbf{x}}_{k+1} = \hat{A}\hat{\mathbf{x}}_k.$$

Similarly for measurements $\mathbf{y}_k$ we have

$$\mathbf{y}_k = C\mathbf{x}_k$$

$$\mathbf{y}_k = CT\hat{\mathbf{x}}_k$$

$$\mathbf{y}_k = \hat{C}\hat{\mathbf{x}}_k.$$

Thus, any given discrete time linear dynamical system can be decomposed into observable and unobservable subsystems. (A more detailed proof can be found in [26].) □

1.2.3 Observable Subspaces

The observable and unobservable subsystems of a system were discussed in Section 1.2.2. A similar approach can be used to separate the state space $\mathbb{R}^n$ into observable and unobservable subspaces.

Assume that the observability matrix Φ in (1.2.5) is not full rank. Then the non-negative definite matrix $\Phi^T\Phi$ has a non-trivial nullspace. In other words, there exists $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\Phi^T\Phi\hat{\mathbf{x}} = 0 \Rightarrow \hat{\mathbf{x}}^T\Phi^T\Phi\hat{\mathbf{x}} = \|\Phi\hat{\mathbf{x}}\|^2 = 0.$$

Thus, for any initial state $\mathbf{x}_0 \in \mathrm{Ker}(\Phi^T\Phi)$ the outputs $\mathbf{y}_k = 0$ and we cannot distinguish the initial condition from the zero vector.

Definition 1.2.4. $\mathcal{N} = \mathrm{Ker}(\Phi^T\Phi)$ is called the ***unobservable subspace*** of the system with dimension $n - \mathrm{rank}(\Phi)$.

Definition 1.2.5. $\mathcal{O} = \mathrm{Im}(\Phi^T\Phi)$ is called the ***observable subspace*** of the system with dimension $\mathrm{rank}(\Phi)$.

Hence, we can decompose $\mathbb{R}^n$ in the following way:

$$\mathbb{R}^n = \mathcal{N} \oplus \mathcal{O}.$$

As can be observed, the rank of the observability matrix determines the dimension of observable subspace, i.e. the bigger the rank of $\Phi^T\Phi$ the more we can observe.

1.2.4 Observability of Continuous Time Linear Dynamical Systems

Although the focus in this dissertation is primarily on discrete time linear dynamical systems, we would like to include a discussion about the observability of continuous time systems, since observability of discrete and continuous time linear dynamical systems are typically studied together in classical texts. Moreover, we will be able to apply some results related to the continuous time systems to our discussion of observability metrics in Chapter 2. To check observability of a continuous time system its **observability Gramian** is used. Consider the continuous time linear dynamical system

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) \tag{1.2.8a}$$

$$\mathbf{y}(t) = C\mathbf{x}(t), \tag{1.2.8b}$$

where $t \geq 0$, $\mathbf{x}(t)$ is the state, $\mathbf{y}(t)$ is the measurement, A is the system matrix and C is the measurement matrix. For any time $t_1 > t_0 \geq 0$ the observability Gramian W is defined as, [29],

$$W(t_0, t_1) = \int_{t_0}^{t_1} e^{A^T(t-t_0)} C^T C e^{A(t-t_0)} dt. \tag{1.2.9}$$

If the system is stable, the steady state observability Gramian $W = W(t_0, \infty)$ is given by the Lyapunov equation as, [29],

$$A^T W + W A = -C^T C.$$

Theorem 1.2.3. If the observability Gramian W is non-singular then the continuous time linear dynamical system defined in (1.2.8) is ***completely observable***.

Proof. Can be found in [28]. □

In addition, for a continuous time linear dynamical system we can also use the observability matrix Φ defined in (1.2.5) to determine its observability, [28]. Furthermore, similar to the discrete time case, a continuous time system, which is not completely observable, can be decomposed into observable and unobservable subsystems, [26], as explained in Section 1.2.2.

The importance of observability of a linear dynamical system is self evident. If the system is not observable, we cannot determine all the components of the state and hence do not have enough information about how the system propagates. However, even if the system is not completely observable and we cannot determine all the components of the state, we can still recover at least its projection onto a subspace of $\mathbb{R}^n$ as long as the observability matrix does not have rank zero. Therefore, our initial goal is to investigate how to design a measurement matrix C to ascertain an observable system.

Chapter 2

TIME-INVARIANT MEASUREMENT SCHEMES

In order to employ observability in sensor scheduling, we first have to find a way to design our sensors which makes the system completely observable. Once the observability of the system is guaranteed, the second challenge is to decide which metric to use for measuring observability of the system and moreover how this metric can be used for sensor scheduling.

In this chapter, we first present classical results for sensor design, defined by the measurement matrix C in (1.2.3), to ascertain the observability of a system with a time-invariant measurement scheme. We then reformulate the classical results to provide explicit guidelines for constructing sensors to ensure the observability of the system. We also discuss different observability metrics, and then demonstrate that using the condition number of the observability matrix in (1.2.5) provides a meaningful observability metric. Finally, we examine the relation between observability matrices and Vandermonde matrices.

2.1 Constructive Observability

Chapter 1 introduced the idea of observability of a linear dynamical system. We would like our system to be completely observable so that we can always determine all the system states. Hence, given an invertible system matrix A a straightforward question would be how to design a measurement matrix C so that the system is completely observable.

This problem has been studied to some extent, see e.g. [28, 29]. Given the system matrix A in (1.2.3), for systems with scalar measurements (i.e. $\mathbf{y}_k \in \mathbb{R}$ and $C = \mathbf{c}^T$,

where $\mathbf{c} \in \mathbb{R}^n$), there are two main tests, called the Popov-Belevitch-Hautus (PBH) tests, which determine the observability of a system depending on the measurement vector $\mathbf{c}^T$, [28]. Although these tests are helpful in determining the observability of a system, they are not constructive in nature. That is, they can be used for checking whether a sensor makes the system observable or not, but they do not provide explicit directions for designing a sensor which makes the system observable.

In this section, we discuss the PBH tests and then reformulate them to present methods to design a measurement matrix C, given the system matrix A, that guarantees the observability of the system.

2.1.1 Popov-Belevitch-Hautus Tests

Theorem 2.1.1. [PBH Eigenvector Test] *The system* $\left\{A, \mathbf{c}^T\right\}$ *is unobservable if and only if there exists a non-zero vector* $\mathbf{p} \in \mathbb{R}^n$ *such that* $A\mathbf{p} = \lambda\mathbf{p}$ *and* $\mathbf{c}^T\mathbf{p} = 0$*, for some* $\lambda \neq 0$*.*

Proof. First assume that such a $\mathbf{p}$ exists. Now consider $\Phi\mathbf{p}$ such that

$$\Phi\mathbf{p} = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T A \\ \vdots \\ \mathbf{c}^T A^{n-1} \end{bmatrix} \mathbf{p} = \begin{bmatrix} \mathbf{c}^T\mathbf{p} \\ \lambda\mathbf{c}^T\mathbf{p} \\ \vdots \\ \lambda^{n-1}\mathbf{c}^T\mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, Φ is singular and the system is unobservable.

Now assume that $\left\{A, \mathbf{c}^T\right\}$ is unobservable. Then, as discussed in Chapter 1, we separate the system into observable and unobservable parts using a transformation matrix T as

$$\begin{bmatrix} \hat{\mathbf{x}}_{k+1}^{(1)} \\ \hat{\mathbf{x}}_{k+1}^{(2)} \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & 0 \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k}^{(1)} \\ \hat{\mathbf{x}}_{k}^{(2)} \end{bmatrix}$$

$$y_k = \begin{bmatrix} (\hat{c}_1)^T & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{k}^{(1)} \\ \hat{\mathbf{x}}_{k}^{(2)} \end{bmatrix}.$$

Then $\mathbf{p}^T = \begin{bmatrix} 0 & \mathbf{p}_{22}^T \end{bmatrix}$, where $\mathbf{p}_{22}$ is an eigenvector of $\hat{A}_{22}$, satisfies the conditions of the theorem. □

Theorem 2.1.2. [**PBH Rank Test**] *The system* $\left\{A, \mathbf{c}^T\right\}$ *is observable if and only if* $\text{rank} \begin{bmatrix} \mathbf{c}^T \\ s\mathbb{I} - A \end{bmatrix} = n$, *for all* $s \in \mathbb{R}$.

Proof. If $\begin{bmatrix} \mathbf{c}^T \\ s\mathbb{I} - A \end{bmatrix}$ has rank n, then there cannot be a non-zero $\mathbf{p} \in \mathbb{R}^n$ such that

$$\begin{bmatrix} \mathbf{c}^T \\ s\mathbb{I} - A \end{bmatrix} \mathbf{p} = 0,$$

for any $s \in \mathbb{R}$, i.e.

$$\mathbf{c}^T\mathbf{p} = 0 \text{ and } A\mathbf{p} = s\mathbf{p}.$$

Then by Theorem 2.1.1 $\left\{A, \mathbf{c}^T\right\}$ should be observable. The converse follows by reversing the arguments. □

2.1.2 Results on the Eigenstructure of A and the Observability Matrix Φ

Although the PBH tests determine if a system is observable for a particular $\mathbf{c}^T$, the tests are not constructive. Thus below we present results on how to construct a measurement vector (or matrix) depending on the eigenstructure of A so that the

system is completely observable. That is, given the dynamics (matrix A) of the system, we will construct the sensor(s) that yield an observable system.

Consider again the discrete time linear dynamical system from (1.2.3), where $\mathbf{u}_k = 0$, given by

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{2.1.1a}$$

$$\mathbf{y}_k = C\mathbf{x}_k, \tag{2.1.1b}$$

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times n}$, $\mathbf{x}_k \in \mathbb{R}^n$, and $\mathbf{y}_k \in \mathbb{R}^m$. We investigate the relation between the eigenstructure of A and the observability matrix Φ by constructing a suitable C that ensures a completely observable system. For this aim, we consider three cases where A has distinct eigenvalues and repeated eigenvalues. We start with the case where A has n distinct eigenvalues. We say that $\mathbf{c} \in \mathbb{R}^n$ has a non-zero weight along each eigenvector of A^T if

$$\mathbf{c} = \sum_{j=1}^{n} \alpha_j \mathbf{q}_j, \tag{2.1.2}$$

and $\alpha_j \neq 0$, $j = 1, \ldots, n$, where $\mathbf{q}_j$ are the right eigenvectors of A^T.

Theorem 2.1.3. *If A has n distinct non-zero eigenvalues then $\left\{A, \mathbf{c}^T\right\}$ is observable if and only if $\mathbf{c} \in \mathbb{R}^n$ has a non-zero weight along each eigenvector of A^T.*

Proof. First consider the observability matrix Φ

$$\Phi = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T A \\ \vdots \\ \mathbf{c}^T A^{n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{c} & A^T\mathbf{c} & \cdots & \left(A^T\right)^{n-1}\mathbf{c} \end{bmatrix}^T.$$

To prove Φ is full rank, we need to show that $\{\mathbf{c}, A^T\mathbf{c}, \ldots, \left(A^T\right)^{n-1}\mathbf{c}\}$ are linearly independent. Since A has n distinct eigenvalues, A^T has n distinct eigenvalues, $\{\lambda_j\}_{j=1}^n$, and n distinct linearly independent eigenvectors, $\{\mathbf{q}_j\}_{j=1}^n$, constituting a basis for $\mathbb{R}^n$.

Thus $\mathbf{c} \in \mathbb{R}^n$ can be written as in (2.1.2) with $\alpha_j \neq 0, j = 1, \ldots, n$. We now write $\left(A^T\right)^k \mathbf{c}$ as

$$\left(A^T\right)^k \mathbf{c} = \sum_{j=1}^{n} \lambda_j^k \alpha_j \mathbf{q}_j.$$

Now consider the equation for linear independence for $\{\mathbf{c}, A^T\mathbf{c}, \ldots, \left(A^T\right)^{n-1}\mathbf{c}\}$ for some coefficients $\mathbf{d} = \begin{bmatrix} d_1 & \cdots & d_n \end{bmatrix}^T \in \mathbb{R}^n$

$$\sum_{k=0}^{n-1} d_{k+1}(\lambda_1^k \alpha_1 \mathbf{q}_1 + \lambda_2^k \alpha_2 \mathbf{q}_2 + \ldots + \lambda_n^k \alpha_n \mathbf{q}_n) = 0. \tag{2.1.3}$$

Since $\mathbf{q}_k$ are linearly independent, (2.1.3) is equivalent to

$$\sum_{k=0}^{n-1} d_{k+1}\lambda_1^k = \sum_{k=0}^{n-1} d_{k+1}\lambda_2^k = \cdots = \sum_{k=0}^{n-1} d_{k+1}\lambda_n^k = 0. \tag{2.1.4}$$

Hence considering the sums as inner products, (2.1.4) yields that $\mathbf{d}$ should be orthogonal to the vectors

$$\left\{ \begin{bmatrix} 1 \\ \lambda_1 \\ \vdots \\ \lambda_1^{n-1} \end{bmatrix}, \begin{bmatrix} 1 \\ \lambda_2 \\ \vdots \\ \lambda_2^{n-1} \end{bmatrix}, \ldots, \begin{bmatrix} 1 \\ \lambda_n \\ \vdots \\ \lambda_n^{n-1} \end{bmatrix} \right\}, \tag{2.1.5}$$

corresponding to the system $V\mathbf{d} = 0$, where

$$V = V(\lambda) = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}, \tag{2.1.6}$$

is the Vandermonde matrix. Since V is invertible, the only $\mathbf{d}$ that would satisfy (2.1.4) is the zero vector, and we conclude that $\{\mathbf{c}, A^T\mathbf{c}, \ldots, \left(A^T\right)^{n-1}\mathbf{c}\}$ are linearly independent so that Φ is full rank.

Now assume that $\left\{A,\ \mathbf{c}^T\right\}$ is observable. The above steps can be reversed to prove the only if part. □

Remark. Theorem 2.1.3 is similar to Theorem 3 in [11], in which the observability problem is approached in the context of data assimilation for discretized partial differential equations.

Note that if any α_j was zero, we would not be able to cover the whole space $\mathbb{R}^n$ with the span of (2.1.5) and it would be possible to find a non-zero $\mathbf{d}$. Thus any zero α_j results in unobservability in the corresponding eigenvector direction.

Corollary 2.1.1. *Assume only $m < n$ many α_j are non-zero, i.e. $\mathbf{c}$ has zero weight along some eigenvectors of A^T. Then applying A^T to $\mathbf{c}$ successively creates at most m linearly independent vectors $\{\mathbf{c}, A^T\mathbf{c}, \ldots, \left(A^T\right)^{m-1}\mathbf{c}\}$.*

Now we check the case where A^T has a repeated eigenvalue and is non-defective.

Theorem 2.1.4. *If A^T has a repeated eigenvalue with algebraic and geometric multiplicity s, then for $C = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_s \end{bmatrix}^T$, where $\mathbf{c}_1$ has a non-zero weight along each eigenvector of A^T and $\mathbf{c}_2, \ldots, \mathbf{c}_s$ are any distinct vectors such that*

$$\{\mathbf{c}_1,\ A^T\mathbf{c}_1,\ \ldots,\ \left(A^T\right)^{n-1}\mathbf{c}_1,\ \mathbf{c}_2,\ \ldots,\ \mathbf{c}_s\}$$

are linearly independent, the system $\{A, C\}$ is completely observable.

Proof. Without loss of generality assume that λ_1 is the repeated eigenvalue. Then there are s eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$ corresponding to λ_1. Since λ_1 has geometric multiplicity s, $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$ are linearly independent. As in the previous proof, the eigenvectors of A^T again constitute a basis for $\mathbb{R}^n$, and $\mathbf{c}_1 \in \mathbb{R}^n$ can be written as in (2.1.2)

$$\mathbf{c}_1 = \sum_{j=1}^{n} \alpha_j \mathbf{q}_j,$$

where $\alpha_j \neq 0, j = 1, \ldots, n$.

By the same reasoning as in Theorem 2.1.3, when $C = \mathbf{c}_1^T$ it is possible to find a vector $\mathbf{d} \in \mathbb{R}^n$ such that it is orthogonal to the vectors

$$\left\{ \begin{bmatrix} 1 \\ \lambda_1 \\ \vdots \\ \lambda_1^{n-1} \end{bmatrix}, \begin{bmatrix} 1 \\ \lambda_2 \\ \vdots \\ \lambda_2^{n-1} \end{bmatrix}, \ldots, \begin{bmatrix} 1 \\ \lambda_{n-s+1} \\ \vdots \\ \lambda_{n-s+1}^{n-1} \end{bmatrix} \right\}.$$

In this case we can only construct $(n - s) + 1$ linearly independent vectors by applying A^T to $\mathbf{c}_1$, and $\{\mathbf{c}_1, A^T\mathbf{c}_1, \ldots, \left(A^T\right)^{n-1}\mathbf{c}_1\}$ only span a $(n-s)+1$ dimensional subspace of $\mathbb{R}^n$. Thus, in order to get a full rank observability matrix Φ we need $(s-1)$ more linearly independent vectors, which make $\{\mathbf{c}_1, A^T\mathbf{c}_1, \ldots, \left(A^T\right)^{n-1}\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ linearly independent in our measurement matrix C. □

Remark. Theorem 2.1.4 is similar to Theorem 4 in [11].

Intuitively, we cannot have complete observability with only one vector $\mathbf{c}_1$ in this case since we cannot distinguish between the directions corresponding to the eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$. So the extra conditions are needed to ensure an observable system.

Corollary 2.1.2. *If A^T has a repeated eigenvalue and is non-defective, then for any $\mathbf{c}^T$, where $\mathbf{c}$ has a non-zero weight along each eigenvector of A^T, the system $\left\{A, \mathbf{c}^T\right\}$ is unobservable.*

Finally, we look at the case where A^T is defective.

Theorem 2.1.5. *If A^T has a repeated eigenvalue with algebraic multiplicity s and geometric multiplicity 1, then for any $\mathbf{c}^T$, where $\mathbf{c} \in \mathbb{R}^n$ has a non-zero weight along each (generalized) eigenvector of A^T, the system $\left\{A, \mathbf{c}^T\right\}$ is completely observable.*

Proof. Without loss of generality assume that λ_1 is the repeated eigenvalue. Since it has geometric multiplicity 1, it has s corresponding generalized eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$ which are linearly independent and can be formulated as

$$\begin{aligned} A^T\mathbf{q}_1 &= \lambda_1\mathbf{q}_1 \\ \left(A^T - \lambda_1 I\right)\mathbf{q}_2 &= \mathbf{q}_1 \\ &\vdots \\ \left(A^T - \lambda_1 I\right)\mathbf{q}_s &= \mathbf{q}_{s-1}. \end{aligned}$$

Thus, we can write the following equations.

$$\begin{aligned} A^T\mathbf{q}_1 &= \lambda_1\mathbf{q}_1 \\ A^T\mathbf{q}_2 &= \mathbf{q}_1 + \lambda_1\mathbf{q}_2 \\ &\vdots \\ A^T\mathbf{q}_s &= \mathbf{q}_{s-1} + \lambda_1\mathbf{q}_s. \end{aligned}$$

Again, since the generalized eigenvectors of A^T constitute a basis for $\mathbb{R}^n$ we can write $\mathbf{c}$ as in (2.1.2). Then after some algebra, $\left(A^T\right)^k\mathbf{c}$ can be expressed as

$$
\begin{aligned}
&\left(A^T\right)^k \mathbf{c} = \\
&\left(\alpha_1\lambda_1^k + k\alpha_2\lambda_1^{(k-1)} + \frac{1}{2}k(k-1)\alpha_3\lambda_1^{(k-2)} + \ldots + \frac{1}{(s-1)!}\frac{k!}{(k-s+1)!}\alpha_s\lambda_1^{(k-s+1)}\right)\mathbf{q}_1 + \\
&\left(\alpha_2\lambda_1^k + k\alpha_3\lambda_1^{(k-1)} + \frac{1}{2}k(k-1)\alpha_4\lambda_1^{(k-2)} + \ldots + \frac{1}{(s-2)!}\frac{k!}{(k-s+2)!}\alpha_s\lambda_1^{(k-s+2)}\right)\mathbf{q}_2 + \\
&\vdots \\
&\left(\alpha_s\lambda_1^k\right)\mathbf{q}_s + \\
&\left(\alpha_{s+1}\lambda_2^k\right)\mathbf{q}_{s+1} + \\
&\vdots \\
&\left(\alpha_n\lambda_{n-s+1}^k\right)\mathbf{q}_n.
\end{aligned}
$$

The above equation is more concisely expressed as

$$
\left(A^T\right)^k \mathbf{c} = \sum_{j=1}^{s}\left(\sum_{i=0}^{s-1}\binom{k}{i}\lambda_1^{(k-i)}\alpha_{i+j}\right)\mathbf{q}_j + \sum_{j=s+1}^{n}\lambda_{j-s+1}^k\alpha_j\mathbf{q}_j.
$$

Following the same reasoning as in the proof of Theorem 2.1.3, to check the linear independence of $\left\{\mathbf{c}, A^T\mathbf{c}, \ldots, \left(A^T\right)^{n-1}\mathbf{c}\right\}$ we should determine the existence of a non-zero vector $\mathbf{d} = \begin{bmatrix} d_0 & \ldots & d_{n-1} \end{bmatrix}^T$ such that it is orthogonal to

$$
\left\{\begin{bmatrix} 1 \\ \lambda_1 \\ \lambda_1^2 \\ \lambda_1^3 \\ \lambda_1^4 \\ \vdots \\ \lambda_1^{(n-1)} \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ s\lambda_1 \\ \vdots \\ \frac{1(n-1)!}{(s-1)!(n-s)!}\lambda_1^{(n-s)} \end{bmatrix}, \begin{bmatrix} 1 \\ \lambda_2 \\ \lambda_2^2 \\ \lambda_2^3 \\ \lambda_2^4 \\ \vdots \\ \lambda_2^{(n-1)} \end{bmatrix}, \ldots, \begin{bmatrix} 1 \\ \lambda_{n-s+1} \\ \lambda_{n-s+1}^2 \\ \lambda_{n-s+1}^3 \\ \lambda_{n-s+1}^4 \\ \vdots \\ \lambda_{n-s+1}^{(n-1)} \end{bmatrix}\right\}.
$$

By confluent Vandermonde matrix properties, [24], these vectors are linearly independent and span $\mathbb{R}^n$. Thus the only $\mathbf{d}$ which is orthogonal to all the above vectors is the zero vector. Hence, Φ is full rank and the system is completely observable. $\square$

Note that even if we had a non-zero weight only along $\mathbf{q}_s$ for the generalized eigenvectors, we would still get a completely observable system.

Corollary 2.1.3. *Assume A^T is as above, for any $\mathbf{c} = \sum_{j=s}^{n} \alpha_j \mathbf{q}_j$, $\alpha_j \neq 0$, $j = s, \ldots, n$, the system is completely observable.*

Proof. Observe that

$$\left(A^T\right)^k \mathbf{c} = \alpha_s \sum_{i=0}^{k} \binom{k}{i} \lambda_1^{(k-i)} \mathbf{q}_{s-i} + \sum_{j=s+1}^{n} \lambda_{j-s+1}^{k} \alpha_j \mathbf{q}_j,$$

which again yields the same set of vectors that are the rows of confluent Vandermonde matrix. Hence, the system is completely observable. $\square$

2.1.3 Equivalency of PBH Tests and Results in Section 2.1.2

In previous discussion we developed methods to construct a measurement matrix C depending on the eigenstructure of A so that the system $\{A, C\}$ is completely observable. Now we establish an equivalency between our results and PBH tests.

Theorem 2.1.3 and PBH Eigenvector Test

Assume that A has n distinct non-zero eigenvalues and $\left\{A, \mathbf{c}^T\right\}$ is unobservable.

Proposition 2.1.1. *There exists a non-zero vector $\mathbf{p} \in \mathbb{R}^n$ such that $A\mathbf{p} = \lambda \mathbf{p}$ and $\mathbf{c}^T p = 0$ if and only if there exists $\alpha_j = 0$, where $c = \sum_{j=1}^{n} \alpha_j \mathbf{q}_j$.*

Proof. Since $\mathbf{q}_j$ are the left eigenvectors of A we have the following relation between the normalized left eigenvectors $\mathbf{q}_j$ and normalized right eigenvectors $\mathbf{p}_l$ of A:

$$\mathbf{q}_j^T \mathbf{p}_l = \delta_{jl}, \tag{2.1.7}$$

where δ_{jl} is the Kronecker delta

$$\delta_{jl} = \begin{cases} 1 & j = l \\ 0 & j \neq l. \end{cases}$$

First assume $\mathbf{p} = \mathbf{p}_l$ for some l such that $A\mathbf{p}_l = \lambda_l \mathbf{p}_l$ and $\mathbf{c}^T \mathbf{p}_l = 0$. Then,

$$\mathbf{c}^T \mathbf{p} = \left(\sum_{j=1}^{n} \alpha_j \mathbf{q}_j^T \right) \mathbf{p}_l = \alpha_l = 0$$

by equation (2.1.7).

Now let $\alpha_l = 0$. Then for the right eigenvector $\mathbf{p}_l$, we have

$$\mathbf{c}^T \mathbf{p}_l = \left(\sum_{j=1}^{n} \alpha_j \mathbf{q}_j^T \right) \mathbf{p}_l = 0$$

by equation (2.1.7). □

Corollary 2.1.2 and PBH Eigenvector Test

Since PBH tests consider the case where $C = \mathbf{c}^T$, we establish an equivalency between the Corollary 2.1.2 and PBH eigenvector test. To simplify the problem, assume the case where the eigenvalue λ_1 has algebraic and geometric multiplicity 2.

Proposition 2.1.2. *For any* $\mathbf{c}$ *with a non-zero weight along each eigenvector of* A^T*, there exists a non-zero vector* $\mathbf{p} \in \mathbb{R}^n$ *such that* $A\mathbf{p} = \lambda \mathbf{p}$ *and* $\mathbf{c}^T \mathbf{p} = 0$*, and hence* $\left\{ A, \mathbf{c}^T \right\}$ *is unobservable.*

Proof. Write $\mathbf{c}$ again as $\mathbf{c} = \sum_{j=1}^{n} \alpha_j \mathbf{q}_j$, where $\alpha_j \neq 0$, $j = 1, \ldots, n$. Assume $A\mathbf{p}_1 = \lambda_1 \mathbf{p}_1$ and $A\mathbf{p}_2 = \lambda_1 \mathbf{p}_2$. Define $\mathbf{p} = \frac{1}{\alpha_1} \mathbf{p}_1 - \frac{1}{\alpha_2} \mathbf{p}_2$. Clearly $A\mathbf{p} = \lambda_1 \mathbf{p}$. Moreover,

$$\mathbf{c}^T\mathbf{p} = \left(\sum_{j=1}^{n-1} \alpha_j \mathbf{q}_j^T\right)\left(\frac{1}{\alpha_1}\mathbf{p}_1 - \frac{1}{\alpha_2}\mathbf{p}_2\right) = \frac{1}{\alpha_1}\alpha_1 - \frac{1}{\alpha_2}\alpha_2 = 0.$$

□

Corollary 2.1.3 and PBH Eigenvector Test

Assume that A^T has a repeated eigenvalue with algebraic multiplicity s and geometric multiplicity 1.

Proposition 2.1.3. *There exists a non-zero vector* $\mathbf{p} \in \mathbb{R}^n$ *such that* $A\mathbf{p} = \lambda\mathbf{p}$ *and* $\mathbf{c}^T\mathbf{p} = 0$ *if and only if there exists* $\alpha_j = 0$*, where* $\mathbf{c} = \sum_{j=s}^{n} \alpha_j \mathbf{q}_j$*.*

Proof. Since only the right eigenvectors $\mathbf{p}_s, \ldots, \mathbf{p}_{n-s+1}$ of A satisfy $A\mathbf{p} = \lambda\mathbf{p}$, the proof is the same as Proposition 2.1.1. Here note that the relation among the right generalized eigenvectors is in reverse order of the left generalized eigenvectors. $A^T\mathbf{q}_1 = \lambda_1\mathbf{q}_1$ but $A\mathbf{p}_1 = \mathbf{p}_2 + \lambda_1\mathbf{p}_1$. Similarly, $A^T\mathbf{q}_s = \mathbf{q}_{s-1} + \lambda_1\mathbf{q}_s$ but $A\mathbf{p}_s = \lambda_1\mathbf{p}_s$.

□

2.2 Metrics of Observability

We now have techniques to design a measurement matrix C so that the system $\{A, C\}$ is completely observable. Now assume that matrices A and C in (2.1.1) are time-invariant but C can be chosen from a collection $\mathcal{C}$ of possible measurements, represented by different measurement matrices C_i. There may be more than one measurement matrix C_i making the system completely observable. Hence, we would like to have a **metric of observability** so that we decide which measurement matrix best suits our system.

For discrete time systems, we formulate the problem as follows. Suppose that we have a linear dynamical system as given in (2.1.1), and a library of possible

measurement matrices $\mathcal{C} = \{C_i\}$. Each C_i gives us a different observability matrix Φ_i. We would like to find Φ_i which optimizes some criterion. The use of observability to select measurements from a collection of potential sensor configurations has been studied, e.g. [13, 51, 52]. Some of these approaches are described below.

2.2.1 A Brief Overview of Current Metrics of Observability

There have been different approaches to measure the observability of a system. In this section, we give a brief overview of some of the currently used methods. Although the metrics are defined using the observability matrix Φ, they can be equivalently used with the observability Gramian W in (1.2.9) .

1. **Smallest eigenvalues (or singular values), [40]**: If a system is near singular then inversion of $\Phi_i^T\Phi_i$ in (1.2.7) and possible errors introduced by the inversion would be dominated by the smallest eigenvalues. Moreover, as the system propagates in time, smallest eigenvalues decay fastest (Assuming they have modulus less than one). Therefore, it is important to capture information pertaining to the eigenmodes corresponding to these eigenvalues before they vanish. This metric would be most relevant if we want to recover the components of $\mathbf{x}_0$ in all eigenmodes with equal precision. One application would be satellite positioning, [34].

2. **Maximizing the spectral radius, [51]**: Another approach is to find the configuration which would maximize the spectral radius of the observability matrix. This can be formulated as $\max \sigma_{\max}(\Phi_i)$. Since $\|\Phi_i\|_2 = \sigma_{\max}(\Phi_i)$, this metric can be considered as an indicator of the geometric size of Φ_i. Larger values for spectral radius correspond to a "bigger" observability matrix.

3. **Maximizing the trace, [51]:** This approach can be formulated as $\max \mathrm{tr}(\Phi_i) = \sum_{j=1}^{n} \sigma_j(\Phi_i)$, where the trace can be interpreted as the size criterion of the singular values, containing the overall average information obtained via the measurements. Hence, this metric would be most beneficial if the obtained information is desired to be maximized on average, without emphasizing any particular eigenmodes. This metric is similar to the A-optimality criterion in experimental design theory, [1].

4. **Figure of merit, [5]:** Three different criteria are combined into a single metric using the Fisher information matrix (FIM), [48], which can be considered as a scaled observability Gramian. The following criteria are combined:

i. Minimizing the **condition number** of the FIM, which is linked to the rank of the matrix and the difficulty in performing its inversion.

ii. Maximizing the **trace** of the FIM, which measures the global sensitivity of the sensors.

iii. Maximizing the **determinant** of the FIM, since inverse of the determinant measures the overall uncertainty of estimation.

Using these criteria, a combined criterion ***figure of merit*** (FOM) can be defined as follows

$$FOM = -\alpha_1 \log(\kappa\,(FIM)) + \alpha_2 \log(\mathrm{tr}\,(FIM)) + \alpha_3 \log(\det\,(FIM)), \qquad (2.2.1)$$

where α_i is the weight of each corresponding criterion.

Example 2.2.1, **[51],** is used to illustrate the problem of selectable measurements.

Example 2.2.1.

Consider the following continuous time linear dynamical system for state variable $x = x(t)$

$$\dot{x} = Ax = \begin{bmatrix} -1 & 1 & 1.5 \\ 1 & -2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

and possible C_i, (2.1.1b),

$$C_1 = [\ 1 \quad 0 \quad 0\ ],\ C_2 = [\ 0 \quad 1 \quad 0\ ],\ C_3 = [\ 0 \quad 0 \quad 1\ ].$$

The results in Table 2.2.1 correspond to some of the criteria mentioned above for the observability Gramians W_1, W_2, W_3 in (1.2.9) for C_1, C_2, C_3.

	W_1	W_2	W_3
$\sigma_{\min}(W_i)$	0.0008	0.0001	**0.0026**
$\sigma_{\max}(W_i)$	**7.00**	2.82	0.41
$\mathrm{tr}(W_i)$	**7.12**	2.92	0.50

Table 2.2.1: Criteria for C_1, C_2, C_3 (For each criterion the "winning" configuration is in bold)

It is evident from Table 2.2.1 that there is no clear "winning" C_i. However, $\sigma_{\max}(W_i)$ and $\mathrm{tr}(W_i)$ indicate that using C_1 and hence measuring x_1 might provide an optimal strategy.

2.2.2 An Analysis of Current Observability Metrics

We now discuss various metrics of observability in detail.

Maximum Eigenvalue, Trace and Determinant of Inverse Characteristics

In [40], the metric problem is approached theoretically with metrics defined axiomatically. While the analysis is conducted for controllability, with appropriate modifications for A, B, C as described in (1.2.2) we can simply replace the word "controllable" with "observable" and obtain the same results. Below the results are presented in their original context, that is, with regard to the controllability of the system.

Consider the continuous time system

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{2.2.2a}$$

$$\mathbf{y}(t) = C\mathbf{x}(t), \tag{2.2.2b}$$

where A is the system matrix, B the input (control) matrix, C the measurement matrix, $\mathbf{x}(t)$ is the state, $\mathbf{u}(t)$ is the input and $\mathbf{y}(t)$ is the measurement. Similar to (1.2.9), the controllability Gramian $W_c(t_0, t_1)$ for $0 < t_0 < t_1 < \infty$ can be written as

$$W_c(t_0, t_1) = \int_{t_0}^{t_1} e^{A(\tau - t_0)} B B^T e^{A^T(\tau - t_0)} d\tau. \tag{2.2.3}$$

Three candidates for physically meaningful metrics are proposed by measuring the minimum control energy (taking the system from $x(t_0) = x_0$ to $x(t_1) = 0$) which is given by

$$\begin{aligned} G(t_0, t_1; x_0) &= \min_u \int_{t_0}^{t_1} \|u(\tau)\|^2 d\tau \\ &= x_0^T W_c^{-1} x_0. \end{aligned}$$

They are:

(i) Maximum eigenvalue of $W_c^{-1}(t_0, t_1)$

This metric measures the *maximum value* of the minimum control energy over the unit ball $\|x_0\| = 1$, i.e.

$$\max_{\|x_0\|=1} G(t_0, t_1; x_0) = \lambda_{\max}(W_c^{-1}) = \frac{1}{\lambda_{\min}(W_c)}$$

yielding

$$\mu_1 = \lambda_{\min}(W_c). \tag{2.2.4}$$

The system is more controllable the larger μ_1 is.

(ii) Trace of $W_c^{-1}(t_0, t_1)$

This metric measures the *average value* of the minimum control energy over the unit ball $\|x_0\| = 1$, given by

$$\begin{aligned}\bar{G}(t_0, t_1) &= \frac{\int_{\|x_0\|=1} x_0^T W_c^{-1} x_0 dx_0}{\int_{\|x_0\|=1} dx_0} \\ &= \frac{1}{n} \operatorname{tr}(W_c^{-1}).\end{aligned}$$

The corresponding metric is then

$$\mu_2 = \frac{n}{\operatorname{tr}(W_c^{-1})}. \tag{2.2.5}$$

Once again, the system is more controllable the larger μ_2 is.

(iii) Determinant of $W_c^{-1}(t_0, t_1)$

This metric utilizes the fact that the volume of the hyperellipsoid $x_0^T W_c^{-1} x_0 = 1$ is proportional to the square root of $\det(W_c)$, i.e.

$$V = \int\limits_{x_0^T W_c^{-1} x_0 \leq 1} dx_0 = \mathcal{O}\left(\sqrt{\det(W_c)}\right).$$

Thus the third metric is

$$\mu_3 = \det(W_c). \tag{2.2.6}$$

We note that since the matrices A, B, C in (2.2.2) are time-invariant, the metrics can also be defined using the controllability matrix Q_c which is given by

$$Q_c = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}. \tag{2.2.7}$$

Thus the metrics become, [40],

$$\mu_i = \mu_i(Q_c Q_c^T), \quad i = 1, 2, 3. \tag{2.2.8}$$

These three metrics are generalized by defining the following axioms.

Definition 2.2.1. [**Axioms of Metric Quality**] For a symmetric positive (semi-) definite matrix P (W_c in (2.2.3) or $Q_c Q_c^T$ in (2.2.7)) a scalar value $\mu(P)$ is called a metric (measure) of quality if and only if

i. $\mu(P) = 0$ if $\det(P) = 0$,

ii. $\mu(P) > 0$ if $\det(P) > 0$,

iii. $\mu(kP) = k\mu(P)$ for $k \geq 0$, (ensuring homogeneity)

iv. $\mu(P_1) \geq \mu(P_2) + \mu(P_3)$ for $P_1 = P_2 + P_3$ (concavity condition: a system with two controls has to be at least as controllable as the two partitioned systems together)

Remark. In [40], it is demonstrated that the metric μ_3 must be modified as

$$\mu_3 = \sqrt[n]{\det(W_c)} \tag{2.2.9}$$

in order to satisfy the axioms.

The metrics μ_i in (2.2.8) can be embedded in the following general metric definition.

Definition 2.2.2. [**Metric Measure**] m_s is called a metric measure and is given by

$$m_s = m_s(\Lambda(P)) = \left(\sum_{i=1}^{n} \frac{1}{n} \lambda_i^s \right)^{1/s}, \tag{2.2.10}$$

where λ_i are the eigenvalues of P.

Using (2.2.10) and (2.2.4), (2.2.5), (2.2.9), we see that $\mu_1 = m_{-\infty}$, $\mu_2 = m_{-1}$, and $\mu_3 = m_0$. Thus, $m_s(\Lambda(P))$ is a metric (measure) for $s \leq 0$, and

$$\mu_1 \leq \mu_2 \leq \mu_3,$$

moreover

$$m_{s_1}(\Lambda(P)) \leq m_{s_2}(\Lambda(P)) \quad if\ s_1 \leq s_2.$$

Condition Number

Following [12], in [13] the condition number of the observability Gramian W_0 in (1.2.9) is proposed as a metric of observability. Their aim is to minimize the errors due to

the inversion of W_0 during reconstruction. We discuss the use of condition number as a metric of observability in more detail in the Section 2.3.

Spectral Radius and Trace

The metrics introduced in [40] and [13] place strong emphasis on the smallest singular values. The reason for selecting these metrics is that if a system is near singular, inversion of the Gramian and errors introduced by the inversion are dominated by the smallest eigenvalues. In [51], it is observed that in order to monitor the principal changes in the state, different metrics should be used. Two new metrics are introduced here:

(i) Spectral Radius of W_0

$$\rho(W_0) = \sigma_{\max}(W_0).$$

This metric can be interpreted as the geometric size of the Gramian. The system is more observable the larger $\rho(W_0)$ is.

(ii) Trace of W_0

$$\mathrm{tr}(W_0) = \sum_{i=0}^{n} \sigma_i(W_0).$$

This metric can be interpreted as the average measure of estimation performance of the sensors. Again the system is more observable the larger $\mathrm{tr}(W_0)$ is.

2.3 Conditioning of Φ as a Metric of Observability

As discussed, there have been efforts to utilize singular values and the conditioning as a metric of observability. For instance, in [52], different measures of observability regarding $\sigma_{\min}(\Phi)$ and $\kappa(\Phi)$ were used to optimize the sensor locations in a linear

system. We consider the condition number $\kappa(\Phi)$ of the observability matrix as a reasonable metric of observability, since to reconstruct $\mathbf{x}_0$ we have to invert $\Phi^T\Phi$, and $\kappa(\Phi)$ relates to both the stability and the accuracy of reconstruction of $\mathbf{x}_0$. Moreover, a small condition number also diminishes the errors in reconstructing $\mathbf{x}_0$ introduced by measurement noise. Thus, given the system matrix A we would like to construct C so that $\kappa(\Phi)$ is minimized.

We first recall the definition of the condition number.

Definition 2.3.1. The 2-norm ***condition number*** $\kappa(M) = \kappa_2(M)$ of $M \in \mathbb{R}^{m\times n}$ is defined by

$$\kappa(M) = \frac{\sigma_{\max}(M)}{\sigma_{\min}(M)},$$

where σ are the singular values of M. If $m = n$, $\kappa(M)$ can be expressed as

$$\kappa(M) = \|M\|_2\|M^{-1}\|_2,$$

where $\|M\|_2$ is the operator norm of M.

To study $\kappa(\Phi)$ as a metric of observability, we consider scalar measurement systems $\left\{A, \mathbf{c}^T\right\}$. First, we construct an upper bound for $\kappa(\Phi)$ in terms of the eigenvalues λ_j of the system matrix A in (2.1.1) and the weights α_j of the measurement vector $\mathbf{c}^T$ in (2.1.2). This provides insight about how λ_j and α_j affect $\kappa(\Phi)$. Second, we inspect the relation between the observability matrix Φ and the Vandermonde matrix of eigenvalues λ_j of A by finding an optimal set of eigenvalues λ_j and a lower bound for $\kappa(\Phi)$ in case of real eigenvalues. We then study the problem of minimizing the condition number $\kappa(\Phi)$ with respect to the measurement vector $\mathbf{c}^T$. Finally, we present some numerical examples.

2.3.1 Upper Bound for $\kappa(\Phi)$

In previous sections we studied the problem of designing measurement matrices C to guarantee complete observability for a linear dynamical system. Even though full rank of Φ ensures observability, an observability matrix with a high condition number would be not practically useful for design purposes, since the condition number of Φ is critical for the accuracy and stability of computing the initial state from the equation $\Phi\mathbf{x}_0 = \mathbf{y}$ in (1.2.6).

Condition number is a highly useful metric to analyze problem sensitivity to perturbations. Assume that we have a scalar measurement system, i.e. $\Phi \in \mathbb{R}^{n\times n}$, and there is some perturbation (noise) in the measurements so that we have $\hat{\mathbf{y}} = \mathbf{y} + \Delta\mathbf{y}$ as our measurement vector. Then $\Phi\mathbf{x}_0 = \mathbf{y}$ is perturbed to

$$\Phi(\mathbf{x}_0 + \Delta\mathbf{x}_0) = \mathbf{y} + \Delta\mathbf{y}.$$

It is easy to show that, [50], the reconstruction error $\Delta\mathbf{x}_0$ can be bounded above by

$$\|\Delta\mathbf{x}_0\| \leq \|\Phi^{-1}\|\|\Delta\mathbf{y}\|.$$

Similarly, if there is perturbation $\Delta\Phi$ in the observability matrix, i.e. $(\Phi + \Delta\Phi)(\mathbf{x}_0 + \Delta\mathbf{x}_0) = \mathbf{y}$, and if we assume $(\Delta\Phi)(\Delta\mathbf{x}_0)$ is negligible, it is not hard to show that

$$\|\Delta\mathbf{x}_0\| \leq \|\Phi^{-1}\|\|\Delta\Phi\|\|\mathbf{x}_0\|.$$

Then using the definition of the condition number of Φ, $\kappa(\Phi) = \|\Phi\|\|\Phi^{-1}\|$, we get

$$\frac{\|\Delta\mathbf{x}_0\|}{\|\mathbf{x}_0\|} \leq \kappa(\Phi)\frac{\|\Delta\mathbf{y}\|}{\|\mathbf{y}\|},$$

$$\frac{\|\Delta \mathbf{x}_0\|}{\|\mathbf{x}_0\|} \leq \kappa\left(\Phi\right) \frac{\|\Delta \Phi\|}{\|\Phi\|}.$$

Hence it is evident that the condition number provides a valuable upper bound for the relative error in reconstructing $\mathbf{x}_0$.

In this section, we investigate the relation between the eigenvalues λ_j of A, the weights α_j of $\mathbf{c}$, and the condition number $\kappa(\Phi)$ of the observability matrix for systems with scalar measurements $\left\{A, \mathbf{c}^T\right\}$. For the case where A is symmetric and has n distinct non-zero eigenvalues, we bound $\kappa(\Phi)$ from above. To achieve this aim we use the following bound, [21],

$$\kappa(\Phi) \leq \frac{2}{|\det(\Phi)|} \left(\frac{\|\Phi\|_F}{\sqrt{n}}\right)^n, \tag{2.3.1}$$

where $\|\Phi\|_F$ is the Hilbert–Schmidt norm of Φ.

In order to use this upper bound we first write the observability matrix Φ as a product of simpler matrices so that we express the upper bound in (2.3.1) in terms of the eigenvalues λ_j and the weights α_j. The procedure is described below.

Assume A is symmetric and has n non-zero distinct eigenvalues. Then the eigenvectors $\{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ of A constitute an orthonormal basis for $\mathbb{R}^n$. Thus, A is diagonalizable and can be written as

$$A = P \Lambda P^T,$$

where $P = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix}$ is the matrix of the eigenvectors of A and is orthogonal since A is symmetric, and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the matrix of eigenvalues of A.

We now write $\mathbf{c}$ in (2.1.2) as

$$\mathbf{c} = P\alpha,$$

where $\alpha = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix}^T$, $\alpha_j \neq 0$. By Theorem 2.1.3, using $\mathbf{c}^T$ guarantees a completely observable system.

We now express the observability matrix as a product of simpler matrices. For this we need to write $\left(A^T\right)^k \mathbf{c} = A^k \mathbf{c}$ using the matrix of eigenvectors P:

$$\begin{aligned} \left(A^T\right)^k \mathbf{c} &= \left(P\Lambda^k P^T\right)(P\alpha) \\ &= P \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \alpha \\ &= P \begin{bmatrix} \lambda_1^k \alpha_1 \\ \vdots \\ \lambda_n^k \alpha_n \end{bmatrix}. \end{aligned} \tag{2.3.2}$$

Hence using (2.3.2), the observability matrix Φ can be expressed as

$$\begin{aligned} \Phi^T &= P \begin{bmatrix} \alpha_1 & \lambda_1\alpha_1 & \cdots & \lambda_1^{n-1}\alpha_1 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n & \lambda_n\alpha_n & \cdots & \lambda_n^{n-1}\alpha_n \end{bmatrix} \\ &= PDV, \end{aligned} \tag{2.3.3}$$

where V is the Vandermonde matrix in (2.1.6) and $D = \operatorname{diag}(\alpha_1, \dots, \alpha_n)$.

Now that Φ is written in terms of three simpler matrices, we express the upper bound (2.3.1) in terms of λ_j and α_j. To do this, we first write $\|\Phi\|_F^2 = \operatorname{tr}(\Phi^T\Phi)$ in (2.3.1) as

$$\begin{aligned}\|\Phi\|_F^2 &= \alpha_1^2 \sum_{k=0}^{n-1} \left(\lambda_1^2\right)^k + \cdots + \alpha_n^2 \sum_{k=0}^{n-1} \left(\lambda_n^2\right)^k \\ &= \sum_{j=1}^{n} \alpha_j^2 \left(\frac{1-\lambda_j^{2n}}{1-\lambda_j^2}\right), \end{aligned} \tag{2.3.4}$$

and

$$\begin{aligned}|\det(\Phi)| &= |\det(P)\det(D)\det(V)| \\ &= |\prod_{j=1}^{n} \alpha_j \prod_{1\le j<l\le n} (\lambda_j - \lambda_l)|.\end{aligned}$$

Here the second line follows since $\det(P) = \pm 1$ due to the orthogonality of P and $\det(V) = \prod_{1\le j<l\le n}(\lambda_j - \lambda_l)$.

We now bound the condition number $\kappa(\Phi)$ of the observability matrix using (2.3.1)

$$\kappa(\Phi) \le \frac{2}{\left|\prod_{j=1}^{n} \alpha_j \prod_{1\le j<l\le n} (\lambda_j - \lambda_l)\right|} \left(\frac{\sqrt{\sum_{j=1}^{n} \alpha_j^2 \left(\frac{1-\lambda_j^{2n}}{1-\lambda_j^2}\right)}}{\sqrt{n}}\right)^n. \tag{2.3.5}$$

The upper bound in (2.3.5) gives us some insight about the relation between $\kappa(\Phi)$ and λ_j, α_j. For instance, we observe that if the eigenvalues λ_j of A are too close to each other, the condition number might be very large. However, numerical tests show that the upper bound (2.3.5) is very loose for Vandermonde-like matrices, which is in agreement with the observations in [21] that the bound is not tight for matrices whose singular values differ by orders of magnitude, e.g. the Vandermonde matrix. To illustrate the looseness of the upper bound (2.3.5) we investigate the case where the eigenvalues λ_j of the system matrix A are uniformly distributed. Since the bound is loose for any type of Vandermonde-like matrix with real nodes, we chose the equispaced case for our numerical example.

Numerical Example

Figure 2.3.1 compares the upper bound (2.3.5) with the computed condition number $\kappa(\Phi)$ of the observability matrix Φ corresponding to A with equispaced eigenvalues on $[-0.9, 0.9]$, $\lambda_j = -0.9 + \frac{j-1}{n-1}1.8$, $j = 1, \ldots, n$, and measurement vector $\mathbf{c}^T$ with weights $\alpha_j = 1$, $j = 1, \ldots, n$. Since our results do not depend on the eigenvectors $\mathbf{p}_j$, any symmetric matrix A with the same eigenvalues would yield similar results.

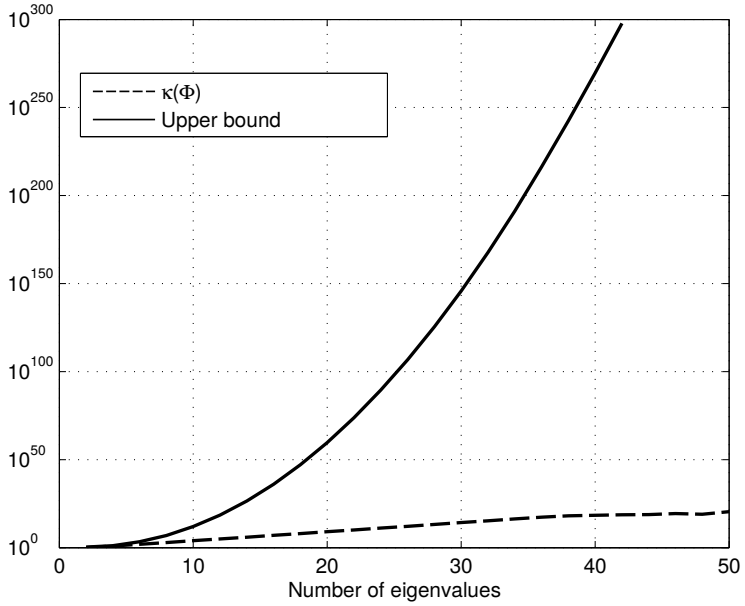

Figure 2.3.1: Comparison of the upper bound (2.3.5) and the computed $\kappa(\Phi)$ for equispaced λ_j

Observe that as the number of eigenvalues (i.e. the dimension of the system) increases, the gap between the computed condition number $\kappa(\Phi)$ and the bound (2.3.5) widens steadily. Still, the bound in (2.3.5) is useful for understanding the relation between the eigenvalues λ_j of the system matrix A, the weights α_j of $\mathbf{c}$ along the eigenvectors and the conditioning of the observability matrix Φ.

2.3.2 Relation between the Observability Matrix and the Vandermonde Matrix

As seen in (2.3.3), the observability matrix Φ of the system $\left\{A, \mathbf{c}^T\right\}$ and V are closely related. Below we show how this relationship can be exploited for some special cases. In particular, we investigate how the eigenvalues λ_j of A can affect $\kappa(\Phi)$.

(i) $\lambda_j = e^{i\frac{2\pi}{n}j}$

Theorem 2.3.1. *If the eigenvalues λ_j of the symmetric system matrix A are roots of unity, i.e. $\lambda_j = e^{i\frac{2\pi}{n}j}$, $j = 1, \ldots, n$ and $\alpha_j = 1$, $j = 1, \ldots, n$, then $\kappa(\Phi) = 1$.*

Proof. By (2.3.3) we can write Φ^T as

$$\Phi^T = PDV.$$

Since $\alpha_j = 1$, we have $D = \mathbb{I}_n$. Moreover, since P is orthogonal and multiplication by an orthogonal matrix does not affect the condition number we get that

$$\kappa(\Phi^T) = \kappa(V).$$

Now, for $\lambda_j = e^{i\frac{2\pi}{n}j}$, V becomes the discrete Fourier transformation matrix, which is known to be unitary. Hence,

$$\kappa(\Phi^T) = \kappa(V) = 1.$$

□

(ii) λ_j **are symmetric on the real axis**

Theorem 2.3.2. *If the eigenvalues λ_j of the symmetric system matrix A are symmetric on real axis and $\alpha_j = 1$, $j = 1, \ldots, n$, then $\kappa(\Phi)$ grows exponentially as the dimension n of the system increases with*

$$\kappa(\Phi) > \frac{2^{n/2}}{n}.$$

Proof. By [17], for λ_j symmetric on real axis, we have

$$\kappa(V) > \frac{2^{n/2}}{n}.$$

By the proof of previous theorem we have $\kappa(\Phi^T) = \kappa(V)$. Thus,

$$\kappa(\Phi^T) > \frac{2^{n/2}}{n}.$$

□

Although Theorem 2.3.2 addresses the case when the eigenvalues are real and symmetric, we observe in practice that these results hold whenever $\lambda_j \in \mathbb{R}$.

2.3.3 Optimal Preconditioner of Vandermonde Matrix

Theorem 2.3.1 illustrates that $\kappa(\Phi)$ is optimal for roots of unity, while Theorem 2.3.2 demonstrates how certain eigenvalues can cause $\kappa(\Phi)$ to grow exponentially. We now ask whether it is possible to minimize $\kappa(\Phi)$ using the weights α_j of $\mathbf{c}$ instead of the eigenvalues λ_j.

For this we recall (2.3.3)

$$\Phi^T = PDV.$$

Since P is orthogonal it does not affect $\kappa(\Phi)$. However we can consider D as a diagonal preconditioner of V. Hence we reformulate the problem as finding the optimal diagonal preconditioner W_{opt} of V such that $\kappa(W_{opt}V)$ is minimized with respect to some norm, i.e.

$$W_{opt} = \arg \min_{W \in \mathbb{R}^{n \times n}} \kappa(WV).$$

The optimal diagonal preconditioner for Vandermonde-like matrices with respect to the Frobenius norm was calculated in [47] as follows:

Define W_{opt} as

$$W_{opt} = \arg \min_{W \in \mathbb{R}^{n \times n}} \kappa_F (WV).$$

If

$$A_j = \|e_j^T V\|_2$$
$$B_j = \| \left(V^T\right)^{-1} e_j \|_2,$$

where $\{e_j\}_{j=1}^n$ is the canonical basis for $\mathbb{R}^n$, then the minimizer W_{opt} is given by

$$W_{opt} = \begin{bmatrix} \sqrt{\frac{B_1}{A_1}} & & & \\ & \sqrt{\frac{B_2}{A_2}} & & \\ & & \ddots & \\ & & & \sqrt{\frac{B_n}{A_n}} \end{bmatrix}. \tag{2.3.6}$$

Hence, by setting $\alpha_j = \sqrt{\frac{B_j}{A_j}}$, $j = 1, \ldots, n$, we obtain the optimal preconditioner for V and thus, the observability matrix Φ has the smallest condition number.

2.3.4 Numerical Tests

We now test our results from Sections 2.3.2 and 2.3.3 with numerical examples.

Growth of $\kappa(\Phi)$

In our first example, we study how fast the condition number $\kappa(\Phi)$ grows for a system with scalar measurements and how it compares to the lower bound in Theorem 2.3.2. Figure 2.3.2 compares the computed condition number $\kappa(\Phi)$ and the lower bound $\frac{2^{n/2}}{n}$ stated in Theorem 2.3.2 for the observability matrix Φ corresponding to a symmetric

system matrix A and measurement vector $\mathbf{c}^T$ with unit weights $\alpha_j = 1, j = 1, \ldots, n$, i.e. $\text{diag}(\alpha_1, \ldots, \alpha_n) = \mathbb{I}_n$. Here we consider A with eigenvalues corresponding to equispaced nodes on $[-0.9, 0.9]$, $\lambda_j^{(1)} = -0.9 + \frac{j-1}{n-1}1.8$, $j = 1, \ldots, n$ and to Chebyshev nodes $\lambda_j^{(2)} = \cos((2j-1)\pi/2n)$, $j = 1, \ldots, n$. Since these two collocation methods are very common in function reconstruction problems, we considered them relevant for our comparison. Note that since the condition number does not depend on eigenvectors, any Φ corresponding to A with the same eigenvalues yield the same results.

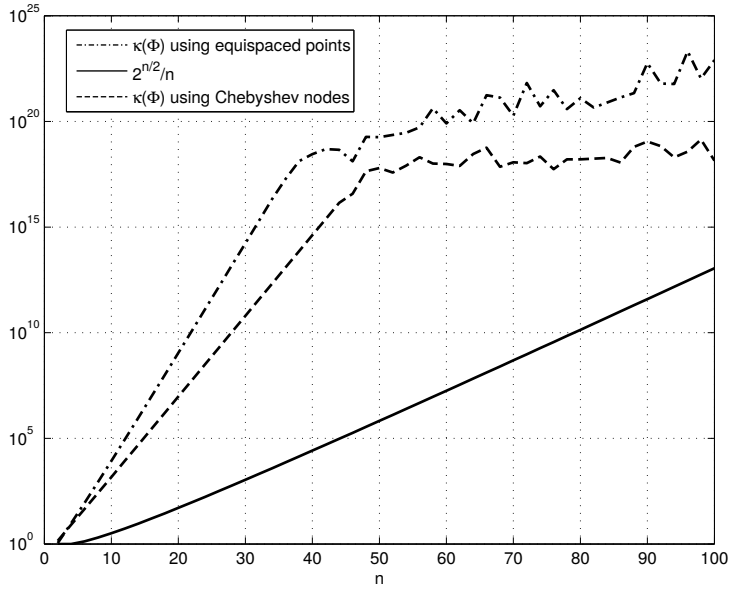

Figure 2.3.2: Comparison of $\kappa(\Phi)$ for $\lambda_j^{(1)}$, $\lambda_j^{(2)}$ and lower bound $\frac{2^{n/2}}{n}$

As can be seen in Figure 2.3.2, $\kappa(\Phi)$ grows exponentially for both equispaced eigenvalues $\lambda_j^{(1)}$ and Chebyshev eigenvalues $\lambda_j^{(2)}$ and is much larger than the lower bound in Theorem 2.3.2. Hence, we observe that $\kappa(\Phi)$ for systems with scalar measurements increases very rapidly as the dimension of the system increases and our accuracy for recovering the initial state $\mathbf{x}_0$ may suffer for high dimensions.

Optimal Preconditioner

In our next example, we would like to investigate how much we can improve $\kappa(\Phi)$ for a system with scalar measurements by using the optimal weights α_j in (2.3.6). Figure 2.3.3 and Figure 2.3.4 compare the computed condition number $\kappa(\Phi)$ of the observability matrix Φ corresponding to measurement vector $\mathbf{c}^T$ with unit weights $\alpha_j^{(1)} = 1,\ j = 1, \ldots, n$ and optimal weights $\alpha_j^{(2)} = \sqrt{\frac{B_j}{A_j}},\ j = 1, \ldots, n$. Here again we consider the cases where the symmetric system matrix A has eigenvalues corresponding to equispaced nodes on $[-0.9, 0.9]$, $\lambda_j^{(1)} = -0.9 + \frac{j-1}{n-1}1.8,\ j = 1, \ldots, n$ and to Chebyshev nodes $\lambda_j^{(2)} = \cos((2j-1)\pi/2n),\ j = 1, \ldots, n$. Figure 2.3.3 compares the condition number $\kappa(\Phi)$ for $\alpha_j^{(1)}$ and $\alpha_j^{(2)}$ for equispaced eigenvalues, and Figure 2.3.4 compares $\kappa(\Phi)$ for $\alpha_j^{(1)}$ and $\alpha_j^{(2)}$ for Chebyshev eigenvalues.

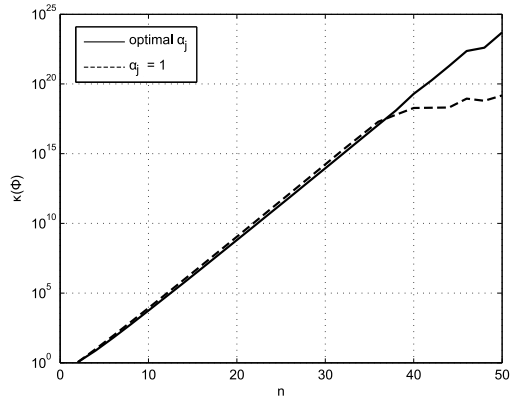

Figure 2.3.3: $\kappa(\Phi)$ for equispaced nodes $\lambda_j^{(1)}$ with $\alpha_j^{(1)} = 1$ and $\alpha_j^{(2)} = \sqrt{\frac{B_j}{A_j}}$

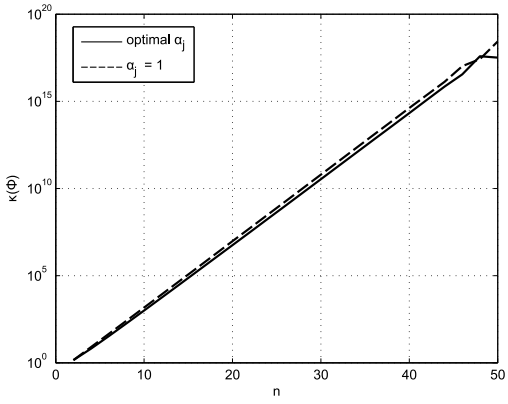

Figure 2.3.4: $\kappa(\Phi)$ for Chebyshev nodes $\lambda_j^{(2)}$ with $\alpha_j^{(1)} = 1$ and $\alpha_j^{(2)} = \sqrt{\frac{B_j}{A_j}}$

As can be seen in the Figures 2.3.3 and 2.3.4, the optimal weights $\alpha_j^{(2)}$ does improve $\kappa(\Phi)$ somewhat about up to $n = 40$. Since the process for finding $\alpha_j^{(2)}$ includes inverting the Vandermonde matrix, possibly for large values of n the accuracy of inversion is low, and we observe that using the unit weights $\alpha_j^{(1)}$ yields better results.

Thus, we conclude that using the optimal weights does not improve the conditioning in a sensible manner.

By the numerical experiments above we see that for systems with scalar measurements the condition number $\kappa(\Phi)$ of the observability matrix grows exponentially and hence, the accuracy of recovering $\mathbf{x}_0$ might be low for high dimensional systems. In addition, we cannot improve $\kappa(\Phi)$ much by employing the weights α_j only.

Chapter 3

TIME-VARIANT MEASUREMENT SCHEMES

3.1 Introduction

In Chapter 2, we considered the situation in which the measurement matrix C in (1.2.3) was invariant for all time steps. Given the system matrix A in (1.2.3), we used the conditioning of the observability matrix as a metric to select the best C from a collection of possible measurement matrices. A natural extension of this idea would be to choose an optimal measurement matrix C_k at each time step k as the system is propagating.

To investigate this problem, we consider systems with time-variant measurement vectors, i.e. $C_k = \mathbf{c}_k^T \in \mathbb{R}^n$. Hence we study the system, $k = 0, 1, 2, \ldots$

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{3.1.1a}$$

$$y_k = \mathbf{c}_k^T \mathbf{x}_k, \tag{3.1.1b}$$

where $\mathbf{x}_k$, $\mathbf{c}_k \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, and $y_k \in \mathbb{R}$.

We let the measurement vector $\mathbf{c}_k$ change over time and select a $\mathbf{c}_k$ at each time step k from a library of possible measurements (sensors) $\mathcal{S} = \{\mathbf{s}_1, \cdots, \mathbf{s}_m\}$, where each $\mathbf{s}_i$ represents a different measurement vector. The collection of measurement vectors $\mathbf{c}_k$ in (3.1.1) over time is called a schedule.

Definition 3.1.1. [(Sensor) **Schedule**] $\{\mathbf{c}_k\}_{k=t_0}^{t_f}$ is called a ***schedule*** for the system (3.1.1) from time step t_0 to t_f, if $\mathbf{c}_{t_0}$ is used as the measurement vector at time t_0, $\mathbf{c}_{t_0+1}$ at time $t_0 + 1$ and so on until $\mathbf{c}_{t_f}$ is used at time t_f.

Since the measurement vector is time-variant in (3.1.1), we need to generalize to definition of the observability matrix.

Definition 3.1.2. [Observability Matrix for Time-Variant Schemes] The observability matrix $\Phi = \Phi(A, \mathbf{c}_k)$ for the system (3.1.1) (from time $k = 0$ to $n-1$) is defined as

$$\Phi = \begin{bmatrix} \mathbf{c}_0^T \\ \mathbf{c}_1^T A \\ \vdots \\ \mathbf{c}_{n-1}^T A^{n-1} \end{bmatrix}. \tag{3.1.2}$$

As discussed in Chapter 2, singular values and the conditioning have been used as a metric of observability for time-invariant measurement systems. For example, in [52] the minimum singular value and condition numbers are used as metrics of observability to optimize the sensor locations. Since reconstructing the initial state $\mathbf{x}_0$ means solving the system $\Phi\mathbf{x}_0 = \mathbf{y}$ in (1.2.6), it is reasonable to consider the condition number $\kappa(\Phi)$ of the observability matrix as a metric of observability. In addition, if white noise is present in measurements, i.e. if the measurements are in the form

$$\mathbf{y}_k = \mathbf{c}_k^T \mathbf{x}_k + \nu_k,$$

where $\nu_k \sim \mathcal{N}[0, \sigma^2]$ for some small variance σ^2, a small condition number helps alleviating the effects of the noise. This follows by the definition of the condition number, since the relative error for reconstructing $\mathbf{x}_0$ is linearly dependent on $\kappa(\Phi)$ and the perturbation, ν_k.

The approach to determining a schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ is to minimize the condition number $\kappa(\Phi)$ corresponding to the system $\left\{A, \mathbf{c}_k^T\right\}_{k=0}^{n-1}$. In this chapter, we first

prove the existence of a schedule that ensures an observable system. We then briefly discuss and present some results for designable measurements. Thereafter we study column subset selection and present a new sensor scheduling algorithm. Finally, as an application of the algorithm, we investigate a diffusive system and present some numerical results.

3.2 Existence of a Schedule

Theorem 3.2.1 shows that if the library $\mathcal{S}$ covers all the eigenmodes of A^T then there always exists a schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ that guarantees observability for the system $\{A, \mathbf{c}_k\}_{k=0}^{n-1}$.

Theorem 3.2.1. *Consider the system (3.1.1). Assume that A has n non-zero distinct eigenvalues, $\{\lambda_j\}_{j=1}^{n}$. Let $\mathcal{S} = \{\mathbf{s}_i\}_{i=1}^{m}$ be a library of possible measurement vectors and $\{\mathbf{q}_j\}_{j=1}^{n}$ be the right eigenvectors of A^T. If for each $\mathbf{q}_j$ there exists an $\mathbf{s}_i$ such that $\mathbf{s}_i$ has a non-zero weight along $\mathbf{q}_j$ in the eigenbasis of A^T, then there exists a schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ such that $\{A, \mathbf{c}_k\}_{k=0}^{n-1}$ is observable.*

Proof. Since A^T has n distinct eigenvalues, $\{\mathbf{q}_j\}_{j=1}^{n}$ constitute a basis for $\mathbb{R}^n$. Then each measurement vector $\mathbf{s}_i$ can be expressed in the eigenbasis of A^T as

$$\mathbf{s}_i = Q\mathbf{a}_i,$$

where $Q = \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix}$ is the matrix of eigenvectors of A^T and $\mathbf{a}_i$ is the vector of coefficients for $\mathbf{s}_i$.

Without loss of generality, assume that $\mathbf{s}_1$ has the most non-zero coefficients in the eigenbasis Q with Z_1 entries corresponding to eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_{Z_1}\}$, i.e. $\arg\max_{i\in\{1,\ldots,m\}} \|\mathbf{a}_i\|_0 = \mathbf{a}_1$, where the 0-norm corresponds to the total number of non-zero elements in a vector. For convention we set $\zeta_1 = Z_1$. Now consider the coefficients

$\mathbf{a}_i^{(1)}$ corresponding to the eigenvectors $Q^{(1)} = \{\mathbf{q}_{Z_1+1}, \ldots, \mathbf{q}_n\}$ for the remaining $\mathbf{s}_i$. Suppose $\mathbf{s}_2$ has the most non-zero coefficients in $Q^{(1)}$, i.e. $\arg\max_{i\in\{2,\ldots,m\}} \|\mathbf{a}_i^{(1)}\|_0 = \mathbf{a}_2^{(1)}$, with ζ_2 entries corresponding to $\{\mathbf{q}_{Z_1+1}, \ldots, \mathbf{q}_{Z_2}\}$ where $Z_2 = \zeta_1+\zeta_2$. We continue in this fashion and finally pick $\mathbf{s}_r$, $r \leq m$, so that it has the most non-zero coefficients in $Q^{(r-1)} = \left\{\mathbf{q}_{Z_{r-1}+1}, \ldots, \mathbf{q}_{Z_r}\right\}$ with ζ_r entries where $Z_r = \zeta_1 + \ldots + \zeta_r = n$. Hence, $\{\mathbf{s}_1, \ldots, \mathbf{s}_r\} \subseteq \mathcal{S}$ cover all the eigenmodes of A^T. In other words,

$$\mathbf{s}_1 = Q \begin{bmatrix} a_{1,1} \\ \vdots \\ a_{1,Z_1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{s}_2 = Q \begin{bmatrix} \vdots \\ a_{2,Z_1+1} \\ \vdots \\ a_{2,Z_2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{s}_r = Q \begin{bmatrix} \vdots \\ a_{r,Z_{r-1}+1} \\ \vdots \\ a_{r,Z_r} \end{bmatrix}.$$

Now consider the following schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$:

$$\begin{aligned} \mathbf{c}_0 &= \mathbf{s}_1 \\ \mathbf{c}_1 &= \mathbf{s}_1 \\ \vdots &= \vdots \\ \mathbf{c}_{Z_1-1} &= \mathbf{s}_1 \\ \mathbf{c}_{Z_1} &= \mathbf{s}_2 \\ \vdots &= \vdots \\ \mathbf{c}_{Z_2-1} &= \mathbf{s}_2 \end{aligned}$$

$$\mathbf{c}_{Z_2} = \mathbf{s}_3$$
$$\vdots = \vdots$$
$$\mathbf{c}_{Z_{r-1}} = \mathbf{s}_r$$
$$\vdots = \vdots$$
$$c_{Z_r-1} = \mathbf{s}_r.$$

The schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ results in the following observability matrix

$$\Phi^T = QV_r,$$

where

$$V_r = \begin{bmatrix} a_{1,1} & \cdots & a_{1,1}\lambda_1^{Z_1-1} & & & \vdots & \\ \vdots & \ddots & \vdots & & & & \\ a_{1,Z_1} & \cdots & a_{1,Z_1}\lambda_{Z_1}^{Z_1-1} & & & & \\ 0 & & & \ddots & & \vdots & \\ \vdots & & & a_{r,Z_{r-1}+1}\lambda_{Z_{r-1}+1}^{Z_{r-1}} & \cdots & a_{r,Z_{r-1}+1}\lambda_{Z_{r-1}+1}^{Z_r-1} \\ & & & \vdots & \ddots & \vdots \\ 0 & & & a_{r,Z_r}\lambda_{Z_r}^{Z_{r-1}} & \cdots & a_{r,Z_r}\lambda_{Z_r}^{Z_r-1} \end{bmatrix},$$

consisting of r many blocks of $\zeta_i \times \zeta_i$ Vandermonde-like matrices $V_{i,i}$ on the diagonal, corresponding to different eigenvalues λ_j and coefficients $\mathbf{a}_i$,

$$V_r = \begin{bmatrix} V_{1,1}(\mathbf{a}_1; \lambda_1, \ldots, \lambda_{Z_1}) & \vdots & & \vdots \\ 0 & V_{2,2}(\mathbf{a}_2; \lambda_{Z_1+1}, \ldots, \lambda_{Z_2}) & & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & V_{r,r}(\mathbf{a}_r; \lambda_{Z_{r-1}+1}, \ldots \lambda_{Z_r}) \end{bmatrix}.$$

Since each $V_{i,i}$ is full rank, V_r is full rank as well. Thus Φ is full rank and, we conclude that the schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ makes the system observable. □

Note that even if we change the time steps when sensors $\{\mathbf{s}_i\}_{i=1}^r$ are used in the schedule, the system would still be observable. Hence, how many times a particular measurement vector $\mathbf{s}_i$ is selected is more critical for ensuring observability than at what time steps k it is used. In particular, if $\mathbf{s}_i$ helps us recover ζ_i many eigenmodes, it has to appear ζ_i times in the schedule for an observable system. However, the order of the sensors clearly matter when a metric is introduced.

3.3 Designable Measurements

Theorem 3.3.1 and 3.3.2 assume that we can design $\mathbf{c}_k$ without restrictions.

Theorem 3.3.1. *Assume A in (3.1.1) is symmetric and full rank, and let $\mathbf{p}_j$ be the normalized eigenvectors of A and λ_j the corresponding eigenvalues. Now let, for $k = 0, \ldots, n-1$,*

$$\mathbf{c}_k = \frac{1}{(\lambda_{k+1})^k}\mathbf{p}_{k+1}.$$

Then for the system $\{A, \mathbf{c}_k^T\}_{k=0}^{n-1}$, $\kappa(\Phi) = 1$.

Proof. Consider the observability matrix Φ

$$\begin{aligned}
\Phi^T &= \begin{bmatrix} \mathbf{c}_0 & A\mathbf{c}_1 & A^2\mathbf{c}_2 & \cdots & A^{n-1}\mathbf{c}_{n-1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{p}_1 & \frac{1}{\lambda_2}A\mathbf{p}_2 & \frac{1}{\lambda_3^2}A^2\mathbf{p}_3 & \cdots & \frac{1}{\lambda_n^{n-1}}A^{n-1}\mathbf{p}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \cdots & \mathbf{p}_{n-1} \end{bmatrix}.
\end{aligned}$$

Since A is symmetric, $\mathbf{p}_i$ are orthonormal. Hence, $\kappa(\Phi) = 1$. □

Thus for the simple case where A is symmetric, assuming $\mathbf{c}_k$ can be designed without restrictions, we can find a schedule $\{\mathbf{c}_k\}$ which minimizes $\kappa(\Phi)$. However, it

should be noted that if A has eigenvalues near zero, the reciprocal terms in $\mathbf{c}_k$ result in numerical instability and the condition number might in fact diverge for these cases.

Theorem 3.3.2 provides an alternative schedule for the case where A does not need to be symmetric.

Theorem 3.3.2. *Assume A in (3.1.1) is invertible and $\{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_{n-1}\}$ is a set of orthonormal vectors in $\mathbb{R}^n$. Let*

$$\mathbf{c}_k^T = \mathbf{r}_k^T (A^k)^{-1}.$$

Then for the system $\left\{A, \mathbf{c}_k^T\right\}_{k=0}^{n-1}$, $\kappa(\Phi) = 1$.

Proof. Consider the measurement y_k of the system state $\mathbf{x}_k = A^k \mathbf{x}_0$

$$\begin{aligned} y_k &= \mathbf{c}_k^T \mathbf{x}_k \\ &= \mathbf{c}_k^T A^k \mathbf{x}_0 \\ &= \mathbf{r}_k^T (A^k)^{-1} A^k \mathbf{x}_0 \\ &= \mathbf{r}_k^T \mathbf{x}_0. \end{aligned}$$

Then the observability matrix Φ becomes

$$\Phi = \begin{bmatrix} \mathbf{r}_0^T \\ \vdots \\ \mathbf{r}_{n-1}^T \end{bmatrix}.$$

Since $\{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_{n-1}\}$ is orthonormal, $\kappa(\Phi) = 1$. □

Thus we see that as long as there are no constraints on the measurement vectors $\mathbf{c}_k$, an optimal schedule such that $\kappa(\Phi) = 1$ can be readily designed.

3.4 Sensor Scheduling

Now assume that instead of being able to design our measurement vectors, we have a library of sensors $\mathcal{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_m\}$ from which we can choose our measurement vector $\mathbf{c}_k$ at each time step k. Then our aim is to choose the best possible sensor $\mathbf{s}_i$ at each time step so that we minimize $\kappa(\Phi)$. However we cannot immediately choose the best conditioned subset of $\mathcal{S}$, since a sensor $\mathbf{s}_i$ used at time k, i.e. $\mathbf{c}_k = \mathbf{s}_i$, results in the row $\mathbf{s}_i^T A^k$ of the observability matrix Φ.

To approach this problem, we first create the $mn \times n$ matrix $M = \Phi^T(A, \mathcal{S})$ as though we are using all the possible sensors in $\mathcal{S}$:

$$M = \left[\begin{array}{c|c|c|c} \mathcal{S} & A^T\mathcal{S} & \cdots & \left(A^T\right)^{n-1}\mathcal{S} \end{array} \right], \tag{3.4.1}$$

where $\left(A^T\right)^k \mathcal{S} = \left(A^T\right)^k \left[\begin{array}{ccc} s_1 & \cdots & s_m \end{array} \right]$, $k = 0, \ldots, n-1$.

We now seek the best conditioned $n \times n$ submatrix $\Phi^T = \left[\begin{array}{ccc} \phi_0 & \cdots & \phi_{n-1} \end{array} \right]$ of M such that the columns of Φ^T satisfy $\phi_k \in \left(A^T\right)^k \mathcal{S}$, $k = 0, \ldots, n-1$. This requirement is consistent with (3.1.1), since only one sensor can be used at each time step. Each column $\phi_k = \left(A^T\right)^k \mathbf{s}_i$, for some $i \in \{1, \ldots, m\}$, of Φ^T corresponds to a different possible sensor, $\mathbf{s}_i$. We can generate our schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ by using the fact that Φ corresponds to the observability matrix of $\{A, \mathbf{c}_k\}_{k=0}^{n-1}$, where

$$\phi_k = \left(A^T\right)^k \mathbf{s}_i = \left(A^T\right)^k \mathbf{c}_k. \tag{3.4.2}$$

3.4.1 Column Subset Selection

The problem of finding the best conditioned submatrix of a given matrix is called column subset selection problem, [7, 8], and can be formulated as follows: Given a

matrix A with n columns and an integer $k < n$, we wish to determine a permutation matrix P such that

$$AP = \begin{bmatrix} A_1 & A_2 \end{bmatrix},$$

where A_1 has the best conditioned k columns of A, and A_2 has the remaining $n-k$ columns. The desirable conditions for subset selection, [8], can be formulated as

i. The smallest singular value $\sigma_k(A_1)$ of A_1 should be as large as possible.

ii. The best linear combination of A_1 should be close to A_2, i.e. $\min_E \|A_1E - A_2\|$, where E is an elementary column operator, should be as small as possible. (A_2 should be well represented by A_1)

3.4.2 RRQR Factorization

One known solution for this problem, when k is the numerical rank of A, is achieved using Rank Revealing QR (RRQR) factorization, [10, 20, 23]. Given an $m \times n$ matrix A with $n \geq m$, RRQR factorization gives the permutation matrix P that yields the QR factorization,

$$AP = QR = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}.$$

The numerical rank k of A can then be determined as follows [20]: R_{11} is well conditioned, $\|R_{22}\|_2$ is small and R_{12} is linearly dependent on R_{11}.
The following definitions and lemmas are needed to perform RRQR factorization.

Definition 3.4.1. An $m \times n$ matrix A has numerical rank k if

$$\sigma_k(A) \gg \sigma_{k+1}(A) = \mathcal{O}(\varepsilon_{mach}),$$

where $\varepsilon_{\text{mach}}$ is the machine precision.

Lemma 3.4.1. [20] *For any permutation matrix P, by the interlacing property of singular values, [18], we have*

$$\sigma_i(R_{11}) \leq \sigma_i(A) \quad \textit{and} \quad \sigma_j(R_{22}) \geq \sigma_{k+j}(A)$$

for $1 \leq i \leq k$ and $1 \leq j \leq n-k$. Thus,

$$\sigma_{\min}(R_{11}) \leq \sigma_k(A),$$
$$\|R_{22}\|_2 = \sigma_{\max}(R_{22}) \geq \sigma_{k+1}(A).$$

Assume $\sigma_k(A) \gg \sigma_{k+1}(A) = \mathcal{O}(\varepsilon_{mach})$, i.e. the numerical rank of A is k. Then we seek P so that $\sigma_{\min}(R_{11})$ is maximized and $\sigma_{\max}(R_{22})$ is minimized.

Lemma 3.4.2. [23] *If $A = QR$ is the QR factorization of A with $R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$ and $R_{11} \in \mathbb{R}^{k\times k}$, and if*

$$\sigma_{\min}(R_{11}) \gg \sigma_{\max}(R_{22}) = \|R_{22}\|_2 = \mathcal{O}(\varepsilon_{mach}),$$

then A has numerical rank k.

Proof. Follows from Lemma 3.4.1. □

Now we give a definition of RRQR factorization.

Definition 3.4.2. [23] Assume that a matrix $A \in \mathbb{R}^{m\times n}$ has numerical rank k. If there exists a permutation $P \in \mathbb{R}^{n\times n}$ such that

$$AP = QR = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

where $R_{11} \in \mathbb{R}^{k\times k}$ and

$$\sigma_{\min}(R_{11}) \gg \|R_{22}\|_2 = \mathcal{O}(\varepsilon_{mach}),$$

then $AP = QR$ is called a Rank Revealing QR (RRQR) factorization of A.

RRQR factorization has been utilized for subset selection problem in different contexts, such as rank deficient least squares, subset selection, and matrix approximation problems [8, 10]. Here we find an RRQR factorization of $M = \Phi^T(A, \mathcal{S})$ in (3.4.1) as

$$MP = Q\begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} = \left[\begin{array}{c|c} \Phi^T & Q\begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix} \end{array}\right],$$

yielding the best conditioned columns of M in Φ^T, from which we determine our sensor schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$.

3.4.3 Sensor Scheduling Algorithm

We now present a new sensor scheduling algorithm which utilizes RRQR factorization for finding the schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$.

Algorithm 3.1 Sensor scheduling algorithm using RRQR factorization

i. Create $M = \Phi^T(A, \mathcal{S})$ in (3.4.1).

ii. Find the permutation matrix P using RRQR factorization $MP = QR$ to obtain the best conditioned columns of M.

iii. Using the first n columns of P, generate P_n so that we work only with the n columns $M_n = MP_n$ of M.

iv. Determine which time steps k and sensors $\mathbf{s}_i$ have been used in M_n.

v. If some time steps are repeated in M_n, find the earliest time step k_0 which has been repeated and the first sensor $\mathbf{s}_0$ used in this time step.

vi. Update M by deleting the column $M(k_0, s_0)$ corresponding to the time step k_0 and sensor s_0 in M.

vii. Repeat steps ii-vi until no time steps are repeated in M_n.

viii. Reorder the columns of M_n to obtain $\hat{M}_n = \Phi^T$ where the order of the columns follows the order of time steps.

ix. Obtain the schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ using (3.4.2).

Algorithm 3.1 essentially deletes one column at a time from M so that in the end each column of M_n comes from a different time step. Hence, it finds the best columns from each time block $\left(A^T\right)^k \mathcal{S}$ and creates a schedule $\{\mathbf{c}_k\}_{k=0}^{n-1}$ which minimizes $\kappa(\Phi)$.

The performance of Algorithm 3.1 depends mainly on the routines used for RRQR factorization. Therefore we now discuss several column subset selection algorithms.

3.5 Column Subset Selection Algorithms

In this section, several different algorithms for column subset selection with their computational complexities are discussed and two of them are compared numerically.

3.5.1 RRQR-MEX

RRQR-MEX provides a MATLAB routine *rrqr*, implementing an interface to the FORTRAN RRQR factorization codes (ACM 782), [4]. The routine has been developed in [46], and the complexity of the ACM 782 algorithm is $\mathcal{O}(n^k)$, [7].

3.5.2 Pan Algorithm

This algorithm was developed in [43]. Rather than an RRQR factorization, it creates a Rank Revealing LU (RRLU) factorization of the matrix A. The RRLU factorization for a square matrix can be described as follows.

Theorem 3.5.1. *Let $A \in \mathbb{R}^{n\times n}$, $1 \leq k < n$ and $\sigma_1 \geq \cdots \geq \sigma_k \geq \sigma_{k+1} \geq \cdots \geq \sigma_n \geq 0$ be the singular values of A. Then there exist permutations Q and P such that*

$$Q^T AP = \begin{bmatrix} L_{11} & 0 \\ L_{21} & I_{n-k} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix},$$

where L_{11} is unit lower triangular and U_{11} is upper triangular,

$$\sigma_k \geq \sigma_{\min}(L_{11}U_{11})$$

and

$$\sigma_{k+1} \leq \|U_{22}\|.$$

An algorithm for the factorization is provided (Algorithm 2, [43]). A comparison to other existing RRQR factorization algorithms yields comparable results, and the complexity is $\mathcal{O}(n^3)$, [7].

3.5.3 Pseudoskeleton Approximation

An investigation of how well A can be approximated with a pseudoskeleton approximation is performed in [19].

Definition 3.5.1. Let $A \in \mathbb{R}^{m\times n}$, $n > m$ and $\text{rank}(A) = r$. Then there exists a non-singular $r \times r$ submatrix $\hat{A}$ in A. If $\hat{A}$ lies in rows $i \in \hat{I} \equiv \{i_1, \dots i_r\}$ and in columns $j \in \hat{J} \equiv \{j_1, \dots j_r\}$, i.e. $\hat{A} = A(\hat{I}, \hat{J})$, then

$$A = C\hat{A}^{-1}R$$

is called a *skeleton decomposition* of A, where

$$C = A(I, \hat{J}),\ R = A(\hat{I}, J)$$

$$I \equiv \{1, \dots, m\},\ J \equiv \{1, \dots, n\}.$$

Now, let $\text{rank}(A+E) = r$, where $\|E\| \approx 0$ for some matrix norm. Then for sufficiently small ε

$$\|A - C\hat{A}^{-1}R\|_2 = \mathcal{O}\left(\|A\|_2^2\|\hat{A}^{-1}\|_2^2\varepsilon\right).$$

If we replace $\hat{A}^{-1}$ with a more suitable matrix B, then we can approximate A by the matrix $B = CGR$. Any matrix of the form $B = CGR$ is called a *pseudoskeleton component* of A.

Theorem 3.5.2. *Assume* $A, F \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(A-F) \leq r$, *and* $\|F\|_2 \leq \varepsilon$ *for some* $\varepsilon > 0$. *Then there exists a pseudoskeleton component* CGR *such that*

$$\|A - CGR\|_2 \leq \mathcal{O}\left(\varepsilon\sqrt{r}\left(\sqrt{m}+\sqrt{n}\right)\right).$$

Methods for finding a suitable G utilizing the SVD decomposition of $\Psi = \hat{A} - \hat{F}$, where $\hat{A}$ and $\hat{F}$ denote the $r \times r$ submatrices which occupy the intersections of rows $\hat{I}$ and columns $\hat{J}$ in A and F, have been presented in [19]. In addition, a MATLAB implementation *SkeletonApproximation* of this method has been presented in [6]. The complexity of the implementation is $\mathcal{O}\left(m^2n + n^3\right)$.

3.5.4 CUR Decomposition

Similar to pseudoskeleton approximation, CUR decomposition, [37], approximates A so that A can be described as a product of its actual rows and columns.

Let A_k be the rank-k SVD approximation of A. Then CUR decomposition seeks a low-rank approximation $CUR \approx A$, where C, R contain small number of actual columns and rows of A, respectively, and U a user constructed matrix such that

$$\|A - CUR\|_F \leq \|A - A_k\|_F + \varepsilon\|A\|_F$$

for some $\varepsilon > 0$. However the problem with this decomposition is that it does not return a fixed number of columns and instead returns an *expected number* of columns. Hence, it cannot be used for the sensor scheduling problem.

3.5.5 Numerical Examples

We compare RRQR-MEX and pseudoskeleton approximation by creating random matrices and comparing the condition number of the returned submatrices and run-

times. All of the examples are implemented in MATLAB on a Linux machine with a 3.20 GHz CPU and 8 GB ram.

(i) $m = 80$, $n = m^2$

In this example, 100 random matrices of size 80×80^2 have been created with 80 columns being selected using the RRQR-MEX and SkeletonApproximation. In the first plot the condition numbers of 80×80 submatrices returned by the algorithms are presented. In the second plot, corresponding run-times in seconds are shown.

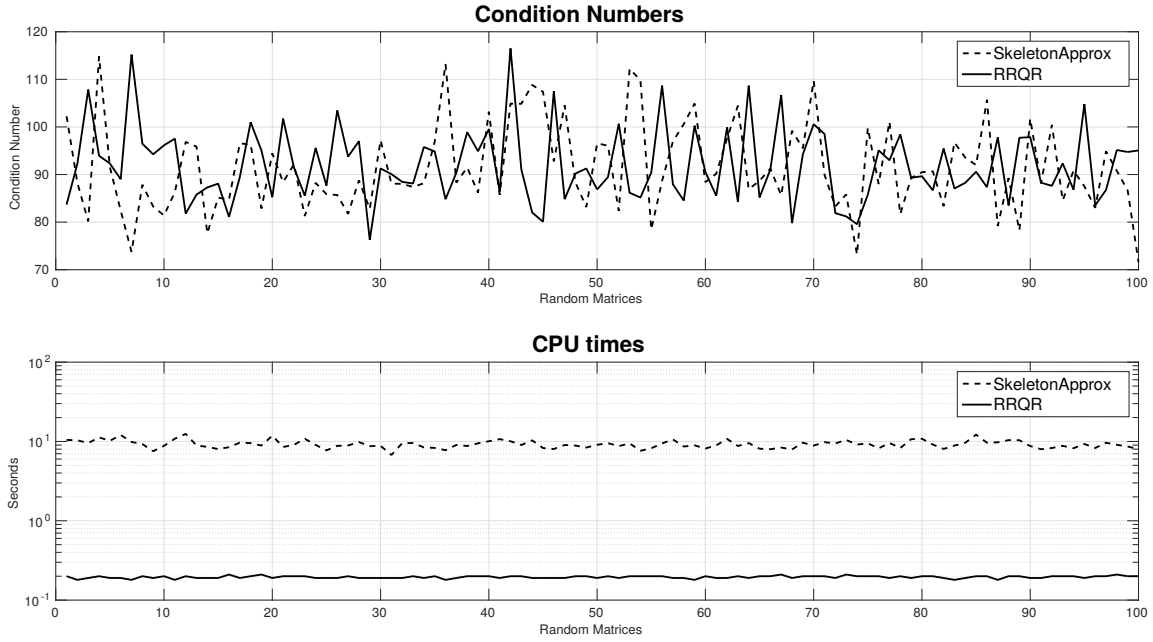

Figure 3.5.1: Condition numbers and run-times for two algorithms

Observe from Figure 3.5.1 that the SkeletonApproximation is nearly 100 times slower. However, since RRQR-MEX is originally coded in FORTRAN and SkeletonApproximation is coded in MATLAB, the run-times cannot be taken as a comparison of their complexities. Rather, we can only conclude that the implementation of pseudoskeleton approximation in MATLAB takes longer than RRQR-MEX's implementation.

(ii) $m = 10, \cdots, 80,\ n = m^2$

In this example, we allow the dimension of the matrices to change. We observe how the condition numbers and the run-times for each algorithm scale. We increase the number of the rows m of the matrix from 10 to 80, and create 100 random matrices of size $m \times m^2$ for each m. Then we select m columns from each random matrix using the algorithms. The averages of the results for 100 matrices have been computed and plotted against m (dimension of the system) in Figure 3.5.2.

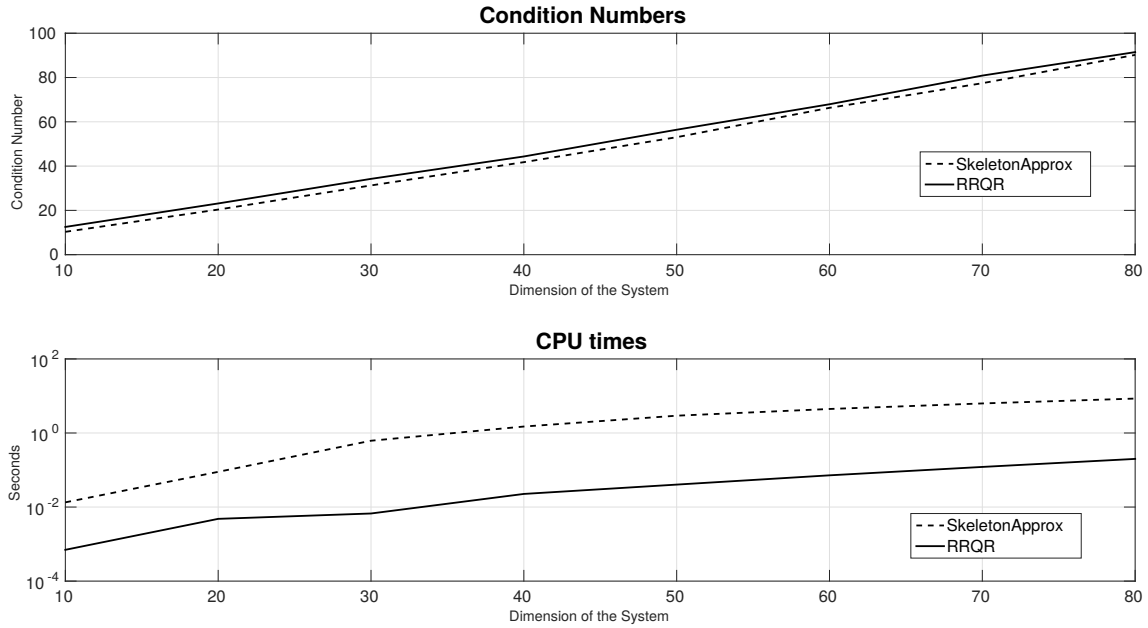

Figure 3.5.2: Average condition numbers and run-times of 100 random matrices for two algorithms as m increases

We can again see that the condition numbers are comparable for both algorithms and SkeletonApproximation is about 100 times slower than RRQR-MEX using MATLAB.

3.6 Diffusion Equation with DST Measurements

We now present a numerical example using diffusion equation to discuss the merits of sensor scheduling using Algorithm 3.1. Consider the $N \times N$ second order centered

difference matrix D_2 for a boundary value problem with $N+2$ collocation points on $[0,1]$

$$D_2 = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix},$$

where

$$h = \frac{1}{N+1}.$$

It is known, [35], that D_2 has the eigenvalues, $p = 1, \dots N$,

$$\lambda_p(D_2) = \frac{2}{h^2}\left(\cos(p\pi h) - 1\right),$$

and eigenvectors

$$\mathbf{u}_p = \begin{bmatrix} \sin(p\pi h) & \cdots & \sin(p\pi N h) \end{bmatrix}^T. \tag{3.6.1}$$

We consider the forward finite difference solution of the diffusion equation

$$u_t = \alpha u_{xx} \tag{3.6.2}$$

for some $\alpha > 0$ on $[0,1]$ with Dirichlet boundary conditions, $u(0) = u(1) = 0$, as our system state. We have the following finite difference matrix as our system matrix A

$$A = \begin{bmatrix} 1-2\gamma & \gamma & 0 & \cdots & 0 \\ \gamma & 1-2\gamma & \gamma & \ddots & \vdots \\ 0 & \lambda & 1-2\gamma & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma \\ 0 & \cdots & 0 & \gamma & 1-2\gamma \end{bmatrix},$$

where

$$\gamma = \alpha \frac{\Delta t}{h^2}.$$

Our system state is the numerical solution $\mathbf{x}_k = \bar{u}(x,t)$ of the heat equation discretized over time and space. For stability, we use

$$\Delta t = rh^2, \tag{3.6.3}$$

with $r < \frac{1}{2}$. For the measurements vectors, we use $\mathbf{u}_p$ in (3.6.1), i.e. the discrete sine transform (DST) library. The library $\mathcal{S} = \{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ of possible measurement vectors can then be written as

$$\mathcal{S} = \left\{ \begin{bmatrix} \sin(k\pi \frac{1}{N+1}) \\ \sin(k\pi \frac{2}{N+1}) \\ \vdots \\ \sin(k\pi \frac{N}{N+1}) \end{bmatrix} \right\}_{k=1}^{N}.$$

Thus we have the scalar measurement system (3.1.1) where the measurement vector $\mathbf{c}_k$ is chosen at each time step k from the DST library S.

3.6.1 Eigenstructure of A and $\Phi^T(A,U)$

We rewrite the system matrix A in terms of D_2 as

$$A = I + \alpha \Delta t D_2.$$

Then A has eigenvalues

$$\begin{aligned} \lambda_p &= 1 + \alpha \Delta t \left(\frac{2}{h^2} \left(\cos(p\pi h) - 1 \right) \right) \\ &= 1 + 2\gamma \left(\cos(p\pi h) - 1 \right), \end{aligned}$$

$p = 1, \ldots N$, and the same eigenvectors $\mathbf{u}_p$ given in (3.6.1).

Defining $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_N \end{bmatrix}$, where $\mathbf{u}_p$ is given in (3.6.1), we can generate $M = \Phi^T(A, U)$ as in (3.4.1)

$$M = \begin{bmatrix} U & A^T U & \cdots & \left(A^T\right)^{N-1} U \end{bmatrix}.$$

Now define the matrix of eigenvalues of D_2 and A, as $\Lambda(D_2)$ and $\Lambda(A)$, respectively. Then

$$D_2 = U\Lambda(D_2)U^{-1} \text{ and } A = U\Lambda(A)U^{-1}.$$

Since U is the DST matrix we have

$$U^T U = \left(\frac{N+1}{2}\right) I \Rightarrow U^{-1} = \left(\frac{2}{N+1}\right) U^T.$$

Thus,

$$\begin{aligned} A^T &= \left(\frac{2}{N+1}\right) U\Lambda(A)U^T \\ A^T U &= \left(\frac{2}{N+1}\right) U\Lambda(A)U^T U \\ &= U\Lambda(A). \end{aligned}$$

Therefore,

$$\begin{aligned} \left(A^T\right)^k U &= U\Lambda^k(A) \\ &= \begin{bmatrix} \lambda_1^k \mathbf{u}_1 & \cdots & \lambda_N^k \mathbf{u}_N \end{bmatrix}. \end{aligned}$$

Thus we see that each time block $\left(A^T\right)^k U$ in M is a column-wise weighted DST matrix U using the powers of the corresponding eigenvalues as weights. Moreover,

for each time block $\left(A^T\right)^k U$ we have the following singular values, $p = 1, \ldots N$, and $k = 0, \ldots, n-1$,

$$\sigma_p\left(\left(A^T\right)^k U\right) = \sqrt{\frac{N+1}{2}}\lambda_p^k.$$

3.6.2 Possible Sensor Schedules and Observability Matrices

For the system to be observable, each $\mathbf{c}_k$ must be a different eigenvector $\mathbf{u}_p$ by the PBH eigenvector test (Theorem 2.1.1). Since the $\mathbf{u}_p$ are orthogonal to each other, all of them are needed to span $\mathbb{R}^n$. Therefore, without loss of generality, any observability matrix Φ has the form

$$\Phi^T = \begin{bmatrix} \lambda_1^{k_1}\mathbf{u}_1 & \cdots & \lambda_N^{k_N}\mathbf{u}_N \end{bmatrix},$$

where $k_i \in \{0, \ldots, N-1\}$, $i = 1, \ldots, N$.

Let $L = \left\{\lambda_1^{k_1}, \ldots, \lambda_N^{k_N}\right\}$. Since the $\mathbf{u}_p$ are orthogonal, the singular values of Φ^T are $\sqrt{\frac{N+1}{2}}L$. Thus, since $\lambda_p < 1$, we can express the condition number $\kappa\left(\Phi^T\right)$ of any observability matrix Φ as

$$\kappa\left(\Phi^T\right) = \frac{\max L}{\min L} = \frac{1}{\min L}.$$

Algorithm 3.1 for the diffusion problem always picks up the higher frequencies first before going down sequentially. This makes intuitive sense, since large frequencies correspond to small eigenvalues, and they decay quickly as time progresses. If we construct the observability matrix using Algorithm 3.1 we get

$$\Phi_{rrqr}^T = \begin{bmatrix} \mathbf{u}_N & \lambda_{N-1}\mathbf{u}_{N-1} & \cdots & \lambda_1^{N-1}\mathbf{u}_1 \end{bmatrix},$$

which yields the singular values, $p = 1, \ldots N$,

$$\sigma_p\left(\Phi_{rrqr}^T\right) = \left(\sqrt{\frac{N+1}{2}}\right)\lambda_{N+1-p}^{p-1}.$$

Hence the condition number for Φ_{rrqr} is

$$\kappa\left(\Phi_{rrqr}^T\right) = \frac{1}{|\min\{\lambda_{N-1}, \lambda_{N-2}^2, \lambda_1^{N-1}\}|}.$$

While it is difficult to prove that the schedule Algorithm 3.1 returns is optimal for large N, an exhaustive search up to dimension $N = 10$ demonstrated that the schedule given by Algorithm 3.1 was in fact optimal in every case. We now compare sequential sampling with our Algorithm 3.1.

3.6.3 Comparison of Sequential Sampling and Algorithm 3.1

Figures 3.6.1 and 3.6.2 compare the results of Algorithm 3.1 to sequential sampling, that is $\mathbf{c}_{k-1} = \mathbf{u}_k$, $k = 1, \ldots, N$, for the diffusion equation with the initial state given by $u_0(x) = x(1-x)$. We chose the parameters $N = 80$, $\alpha = 0.1$ in (3.6.2) and $r = 0.4$ in (3.6.3). In Figure 3.6.1, the first plot shows the sensors ($\mathbf{u}_p$) chosen by Algorithm 3.1, and the second plot displays the measurements $y_k = \mathbf{c}_k^T\mathbf{x}_k$ corresponding to sequential schedule and Algorithm 3.1 (RRQR) schedule.

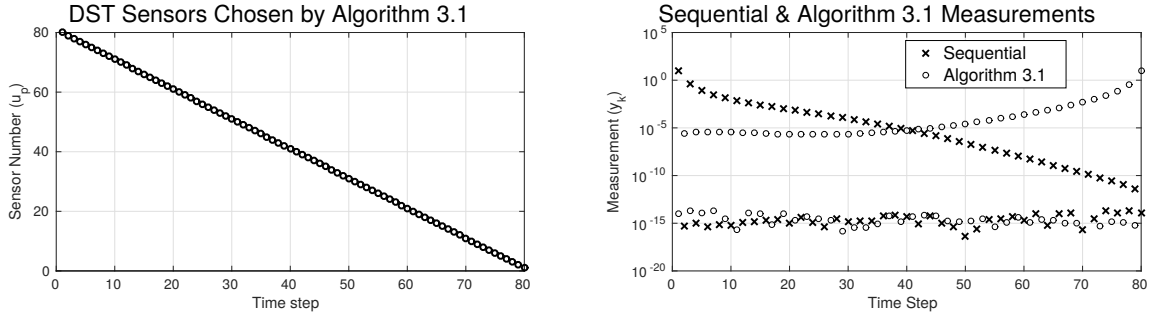

Figure 3.6.1: Sensors chosen by Algorithm 3.1 and measurements for sequential and Algorithm 3.1 schedules

Figure 3.6.2 compares the pointwise reconstruction errors of $\mathbf{x}_0$ for both schedules. The corresponding condition numbers of observability matrices and the two norm

reconstruction errors are also included in the plot legends. As we observe, Algorithm 3.1 yields a five orders of magnitude improvement in the condition number, and a four orders of magnitude improvement in the reconstruction. We note that numerical tests verify that using a random sampling instead of sequential sampling generally yields similar results.

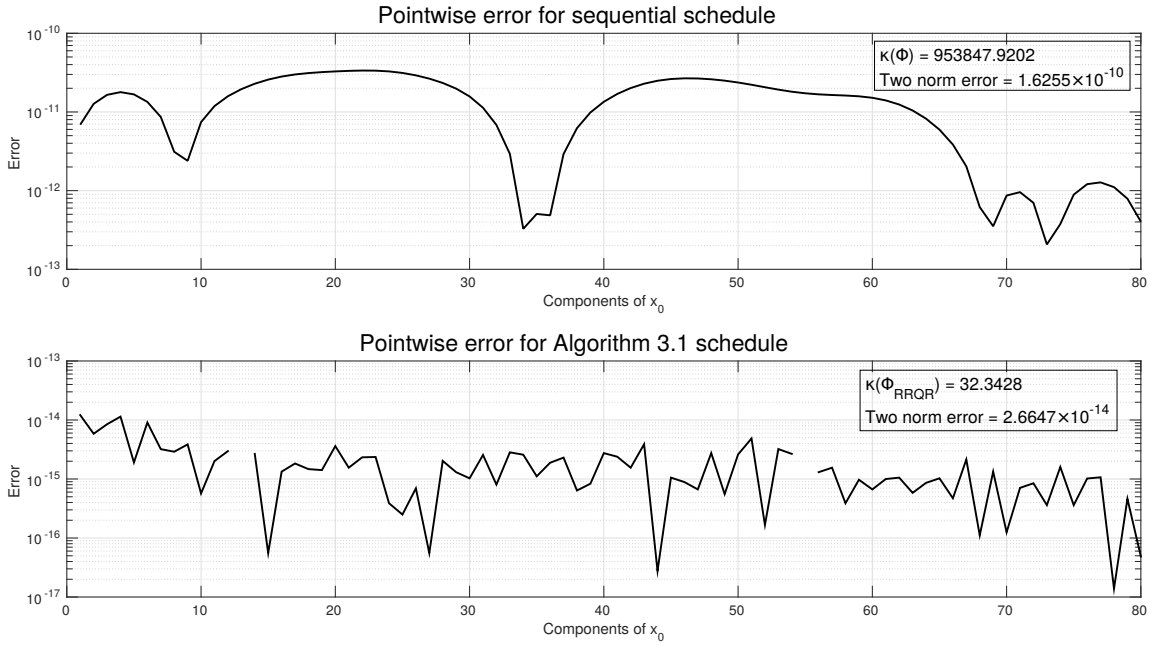

Figure 3.6.2: Reconstruction errors for sequential and Algorithm 3.1 schedules

3.6.4 Noisy Measurements

While either scheduling algorithm is sufficiently robust in the absence of noise, the condition number matters when noise is present, since the perturbation in the measurements is amplified by the condition number. We demonstrate this below.

Assume that the measurements y_k are corrupted by additive white noise so that

$$y_k = \mathbf{c}_k^T \mathbf{x}_k + \nu_k,$$

where $\nu_k \sim \mathcal{N}[0, \sigma^2]$. To analyze Algorithm 3.1 in noisy environments, we must consider the relation between relative error, condition number and the signal-to-noise ratio (SNR).

Theorem 3.6.1. *The relative error in reconstructing* $\mathbf{x}_0$ *is linearly dependent on the condition number of* Φ *and the inverse of SNR.*

Proof. For the noiseless case we have

$$\Phi \mathbf{x}_0 = \mathbf{y},$$

and in the presence of the noise we have

$$\Phi \hat{\mathbf{x}}_0 = \mathbf{y} + \nu,$$

where $\mathbf{y}$ is the vector of measurements and ν is the noise vector. Therefore $\Phi(\mathbf{x}_0 - \hat{\mathbf{x}}_0) = \nu$, and

$$\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\| \leq \|\Phi^{-1}\| \|\nu\|.$$

Now consider the relative error

$$\begin{aligned} \frac{\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|}{\|\mathbf{x}_0\|} &\leq \frac{\|\Phi^{-1}\| \|\nu\|}{\|\mathbf{x}_0\|} \\ &\leq \frac{\|\Phi^{-1}\| \|\nu\|}{\|\mathbf{x}_0\|} \frac{\|\Phi\| \|\mathbf{x}_0\|}{\|\mathbf{y}\|} \\ &\leq \kappa(\Phi) \frac{\|\nu\|}{\|\mathbf{y}\|}. \end{aligned} \tag{3.6.4}$$

□

By (3.6.4), we can clearly see that the relative reconstruction error is dependent on the condition number. Hence, keeping the condition number low helps to reduce the effect of the noise on the reconstruction.

We now show some numerical results regarding the preceding discussion.

3.6.5 Comparison of Sequential Sampling and Algorithm 3.1 with Noisy Measurements

We simulate the noisy case by changing the variance σ^2 of the noise from 10^{-16} to 10^{-1} and observe its effects on the relative error. We also plot the bound in (3.6.4) to see how tight it is.

Figure 3.6.3 compares the results for sequential sampling and Algorithm 3.1 using the same parameters as in the previous example for $N = 40, 80$. We made 100 simulations for each variance value and plotted the average relative errors along with the bounds.

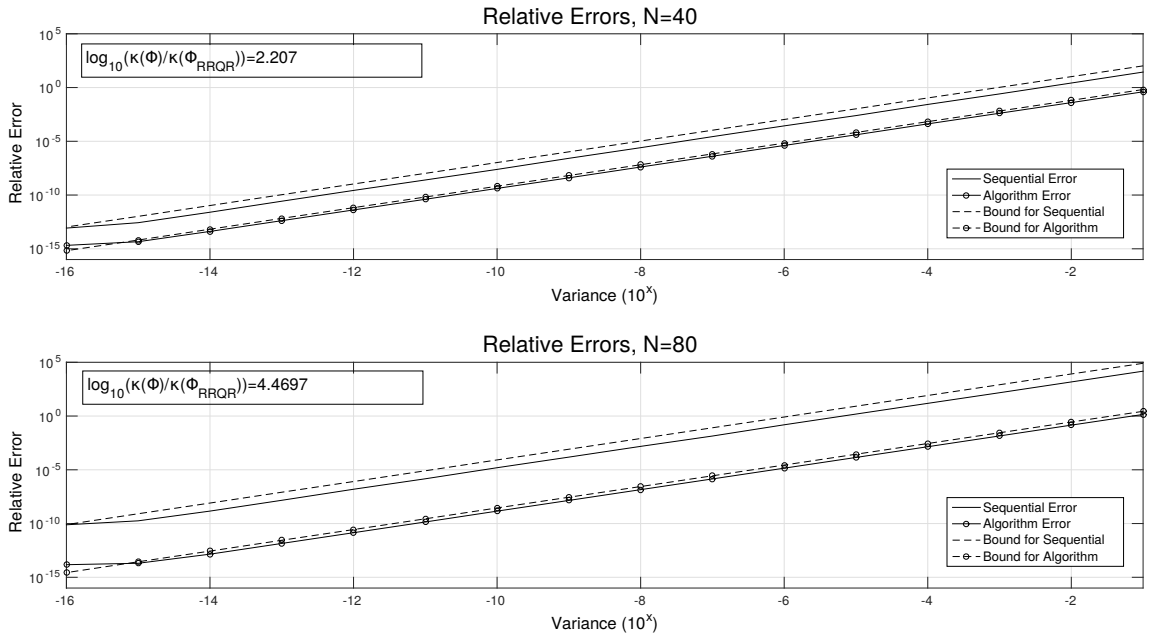

Figure 3.6.3: Relative errors for sequential sampling and Algorithm 3.1 for $N = 40, 80$

Observe that the relative error is improved in the order of $\mathcal{O}(10^2)$ for $N = 40$, and in the order of $\mathcal{O}(10^4)$ for $N = 80$. Hence we see that the ratio of the relative errors follows the ratio of the condition numbers. Moreover, although the error resulting from Algorithm 3.1 stays almost the same for both $N = 40$ and 80, the error for

sequential sampling increases almost three orders of magnitude. We further observe that the bound in (3.6.4) is tight in this example, suggesting that it can be used as a good estimate for the relative error.

Chapter 4

CONCLUSION

This dissertation discussed observability and its multiple uses for linear systems with both time-invariant and time-variant measurement schemes. Specifically, we employed the condition number as a metric of observability, and used this metric for designing measurement schemes and generating sensor schedules.

For time-invariant measurement schemes, we reformulated PBH tests (Theorems 2.1.1 and 2.1.2) to provide guidelines for constructing a measurement vector or matrix that ensures observability. The advantage of this approach is evident for designing measurement schemes. With explicit methods, one can easily construct sensors which would make a system observable, rather than trying to find a working sensor setup from possible sensor configurations. Another advantage of this approach is that since the construction of the measurement vector $\mathbf{c}$ in (2.1.2) depends on the eigenstructure of A, one can easily design a sensor that emphasizes some eigenmodes of A over the others by adjusting the weights α_j in (2.1.2). Hence if some eigenmodes are more critical than others, these eigenmodes can be measured more closely without sacrificing overall observability.

Although the upper bound (2.3.5) for the condition number $\kappa(\Phi)$ of the observability matrix Φ is not tight, it provides insight about how the eigenvalues λ_j of A and the weights α_j of $\mathbf{c}$ affect $\kappa(\Phi)$. As suggested by the bound (2.3.5), λ_j play a bigger role than α_j in determining the magnitude of $\kappa(\Phi)$. Moreover, with our analysis on the similarities between the Vandermonde matrix and Φ, we were able to observe that for scalar measurement systems, adjusting the weights α_j was not very effective for diminishing $\kappa(\Phi)$.

Observability has been utilized for sensor selection [13, 51, 52]; however to our knowledge, it has not been studied much in the context of sensor scheduling. As in the case for time-invariant measurement schemes, recent methods are useful for selecting an optimal sensor configuration from a set of possible configurations with respect to some observability metric. However, they are not practical for sensor scheduling, where a different measurement has to be chosen at each time step rather than selecting an optimal time-invariant measurement matrix. Using the condition number metric, we provided results for designing schedules and presented a new sensor scheduling algorithm using observability and RRQR factorization. One advantage of Algorithm 3.1 is that it can be computed a priori, and hence does not require any online computation time. Moreover, since Algorithm 3.1 only uses the system matrix A in (3.1.1) and the set of possible sensors S for generating a schedule, it can reduce the need for running or simulating the system in order to obtain a schedule.

As an application of Algorithm 3.1 we studied diffusion equation with DST measurements. This analysis can be generalized to other applications where there is diffusion in the system. One such application is MRI measurements, where the magnetic field decays over time. In addition, since MRI measurements are computationally expensive, being able to generate an a priori schedule would greatly assist reducing the computational cost of finding a schedule online.

In this dissertation we focused on scalar measurement systems, i.e. $y_k = \mathbf{c}_k^T \mathbf{x}_k \in \mathbb{R}$ in (3.1.1). One natural continuation of our work would be to study the multi-dimensional measurement case where $\mathbf{y}_k \in \mathbb{R}^m$. Moreover, systems with noisy measurements have been discussed to a short extent in Section 3.6.4. The analysis for noisy measurements can be expanded, and Algorithm 3.1 can be improved in order to better address the case when measurements contain noise. For example, instead of assuming a zero mean noise we can have the measurement noise $\nu_k \sim \mathcal{N}[\mu, \sigma^2]$ for

some $\mu \neq 0$ and can use $\hat{y}_k = y_k - \mu$ to apply our method. In addition, we might assume noise in the sensors, i.e $\mathbf{s}_i = \bar{\mathbf{s}}_i + \mathbf{w}_i$ for some noise parameter $\mathbf{w}_i$, which will result in a random observability matrix. We can also consider the case where the sensors depend on a continuous parameter, i.e. $\mathbf{s}_i = \mathbf{s}_i(\theta)$, for some $\theta \in \mathbb{R}$, such as gain of a sensor. We can utilize this property, for instance, to emphasize the eigenmodes corresponding to small eigenvalues. In this case we would have a continuous observability matrix.

Finally, since controllability is the dual concept of observability, a similar study can be conducted for controllability.

REFERENCES

[1] A. Atkinson and A. Donev, *Optimum experimental designs*, Oxford University Press, 1992.

[2] M. Babaali and M. Egerstedt, *Observability of switched linear systems*, in Hybrid Systems: Computation and Control, Springer, 2004, pp. 48–63.

[3] F. Bian, D. Kempe, and R. Govindan, *Utility based sensor selection*, in Proceedings of the 5th International Conference on Information Processing in Sensor Networks, ACM, 2006, pp. 11–18.

[4] C. H. Bischof and G. Quintana-Ortí, *Algorithm 782: Codes for rank-revealing QR factorizations of dense matrices*, ACM Transactions on Mathematical Software (TOMS), 24 (1998), pp. 254–257.

[5] S. Borguet and O. Léonard, *The Fisher information matrix as a relevant tool for sensor selection in engine health monitoring*, International Journal of Rotating Machinery, 2008 (2008).

[6] C. Boutsidis, *A theory of pseudoskeleton approximations*. `http://www.boutsidis.org/software.html`. Accessed: 2015-06-13.

[7] C. Boutsidis, M. W. Mahoney, and P. Drineas, *An improved approximation algorithm for the column subset selection problem*, in Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2009, pp. 968–977.

[8] M. E. Broadbent, M. Brown, K. Penner, I. Ipsen, and R. Rehman, *Subset selection algorithms: Randomized vs. deterministic*, SIAM Undergraduate Research Online, 3 (2010), pp. 50–71.

[9] A. Carmi, *Sensor scheduling via compressed sensing*, in Proceedings of the 13th Conference on Information Fusion, IEEE, 2010, pp. 1–8.

[10] T. F. Chan and P. C. Hansen, *Some applications of the rank revealing QR factorization*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 727–741.

[11] S. E. Cohn and D. P. Dee, *Observability of discretized partial differential equations*, SIAM Journal on Numerical Analysis, 25 (1988), pp. 586–617.

[12] T. Damak, J. Babary, and M. Nihtila, *Observer design and sensor location in distributed parameter bioreactors*, in IFAC Symposia Series, Pergamon Press, 1993, pp. 87–87.

[13] D. Dochain, N. Tali-Maamar, and J. Babary, *On modelling, monitoring and control of fixed bed bioreactors*, Computers & Chemical Engineering, 21 (1997), pp. 1255–1266.

[14] E. Ertin, J. W. Fisher, and L. C. Potter, *Maximum mutual information principle for dynamic sensor query problems*, in Information Processing in Sensor Networks, Springer, 2003, pp. 405–416.

[15] T. E. Fortmann and K. L. Hitz, *An introduction to linear control systems*, Crc Press, 1977.

[16] Z. Gajic and M. Lelic, *Modern control system engineering*, Prentice Hall International, Intl. Series in Systems and Control Engineering, London, 1996.

[17] W. Gautschi and G. Inglese, *Lower bounds for the condition number of Vandermonde matrices*, Numerische Mathematik, 52 (1987), pp. 241–250.

[18] G. H. Golub and C. F. Van Loan, *Matrix computations*, vol. 3, JHU Press, 2012.

[19] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, *A theory of pseudoskeleton approximations*, Linear Algebra and Its Applications, 261 (1997), pp. 1–21.

[20] M. Gu and S. C. Eisenstat, *Efficient algorithms for computing a strong rank-revealing QR factorization*, SIAM Journal on Scientific Computing, 17 (1996), pp. 848–869.

[21] H. W. Guggenheimer, A. S. Edelman, and C. R. Johnson, *A simple estimate of the condition number of a linear system*, The College Mathematics Journal, 26 (1995), pp. 2–5.

[22] F. J. Hale, *Introduction to control system analysis and design*, Prentice Hall PTR, 1973.

[23] Y. P. Hong and C.-T. Pan, *Rank-revealing QR factorizations and the singular value decomposition*, Mathematics of Computation, 58 (1992), pp. 213–232.

[24] S.-H. Hou and W.-K. Pang, *Inversion of confluent Vandermonde matrices*, Computers & Mathematics with Applications, 43 (2002), pp. 1539–1547.

[25] G. E. Hovland and B. J. McCarragher, *Dynamic sensor selection for robotic systems*, in Proceedings of IEEE International Conference on Robotics and Automation, vol. 1, IEEE, 1997, pp. 272–277.

[26] M. Jamshidi and M. Malek-Zavarei, *Linear control systems: A computer-aided approach*, Butterworth-Heinemann, 1986.

[27] S. Joshi and S. Boyd, *Sensor selection via convex optimization*, IEEE Transactions on Signal Processing, 57 (2009), pp. 451–462.

[28] T. Kailath, *Linear systems*, vol. 1, Prentice-Hall Englewood Cliffs, 1980.

[29] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, vol. 1, Prentice Hall Upper Saddle River, NJ, 2000.

[30] R. Kalman, *On the general theory of control systems*, IRE Transactions on Automatic Control, 4 (1959), pp. 110–110.

[31] R. E. Kalman, *Contributions to the theory of optimal control*, Bol. Soc. Mat. Mexicana, 5 (1960), pp. 102–119.

[32] R. E. Kalman, *Canonical structure of linear dynamical systems*, Proceedings of the National Academy of Sciences, 48 (1962), p. 596.

[33] R. E. Kalman, *Mathematical description of linear dynamical systems*, Journal of the Society for Industrial & Applied Mathematics, Series A: Control, 1 (1963), pp. 152–192.

[34] D. C. Kammer, *Sensor placement for on-orbit modal identification and correlation of large space structures*, Journal of Guidance, Control, and Dynamics, 14 (1991), pp. 251–259.

[35] R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems*, vol. 98, SIAM, 2007.

[36] K. Lim, *Method for optimal actuator and sensor placement for large flexible structures*, Journal of Guidance, Control, and Dynamics, 15 (1992), pp. 49–57.

[37] M. W. Mahoney and P. Drineas, *CUR matrix decompositions for improved data analysis*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 697–702.

[38] A. J. Marques, *On the relative observability of a linear system*, Master's thesis, Naval Postgraduate School Monterey CA, 1986.

[39] R. Monzingo, *A note on sensitivity of system observability*, IEEE Transactions on Automatic Control, 12 (1967), pp. 314–315.

[40] P. Müller and H. Weber, *Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems*, Automatica, 8 (1972), pp. 237–246.

[41] K. R. Muske and C. Georgakis, *Optimal measurement system design for chemical processes*, AIChE Journal, 49 (2003), pp. 1488–1494.

[42] Y. Oshman, *Optimal sensor selection strategy for discrete-time state estimators*, IEEE Transactions on Aerospace and Electronic Systems, 30 (1994), pp. 307–314.

[43] C.-T. Pan, *On the existence and computation of rank-revealing LU factorizations*, Linear Algebra and its Applications, 316 (2000), pp. 199–222.

[44] H. Rowaihy, S. Eswaran, M. Johnson, D. Verma, A. Bar-Noy, T. Brown, and T. La Porta, *A survey of sensor selection schemes in wireless sensor networks*, in Defense and Security Symposium, International Society for Optics and Photonics, 2007, pp. 65621A–65621A.

[45] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards, *Artificial intelligence: A modern approach*, vol. 74, Prentice Hall Englewood Cliffs, 1995.

[46] J. Saak, *RRQR-MEX a MATLAB mex-interface for the rank revealing QR factorization*. `https://www2.mpi-magdeburg.mpg.de/mpcsc/mitarbeiter/saak/Software/rrqr.php?lang=en`. Accessed: 2015-06-13.

[47] J. Saraswat, *A study of Vandermonde-like matrix systems with emphasis on preconditioning and Krylov matrix connection.*, PhD thesis, University of Kansas, 2009.

[48] M. J. Schervish, *Theory of statistics*, Springer, 1995.

[49] D. Sinno and D. Cochran, *Dynamic estimation with selectable linear measurements*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, IEEE, 1998, pp. 2193–2196.

[50] L. N. Trefethen and D. Bau III, *Numerical linear algebra*, no. 50, SIAM, 1997.

[51] F. Van den Berg, H. Hoefsloot, H. Boelens, and A. Smilde, *Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria*, Chemical Engineering Science, 55 (2000), pp. 827–837.

[52] W. Waldraff, D. Dochain, S. Bourrel, and A. Magnus, *On the use of observability measures for sensor location in tubular reactor*, Journal of Process Control, 8 (1998), pp. 497–505.

[53] T.-S. Yoo and S. Lafortune, *NP-completeness of sensor selection problems arising in partially observed discrete-event systems*, IEEE Transactions on Automatic Control, 47 (2002), pp. 1495–1499.

[54] T. K. Yu and J. H. Seinfeld, *Observability and optimal measurement location in linear distributed parameter systems*, International Journal of Control, 18 (1973), pp. 785–799.